Three Essays on Economic Behavior: Evidence from Professional Tennis

Dissertation
submitted to the Faculty of Business, Economics and Informatics of the University of Zurich

to obtain the degree of

Doktor der Wirtschaftswissenschaften, Dr. oec.
(corresponds to Doctor of Philosophy, PhD)

presented by

Paolo Bizzozero
from Vezia, Ticino

approved in October 2017 at the request of

Prof. Dr. Egon Franck
Prof. Dr. Helmut Dietl

The Faculty of Business, Economics and Informatics of the University of Zurich hereby authorizes the printing of this dissertation, without indicating an opinion of the views expressed in the work.

Zurich, October 25, 2017

Chairman of the Doctoral Board: Prof. Dr. Steven Ongena

Contents

Acknowledgments

This dissertation combines almost four years of exciting research at the Chair of Strategic Management and Business Policy at the University of Zurich. Along this journey, there have been a number of important people who have helped me greatly and whom I would like to thank.

First and foremost, I want to sincerely thank Egon Franck. Collaborating with him throughout these years, both in teaching and research, has been a great privilege. Under his supervision, I learned a great deal, and thanks to his patience and enduring support, I was able to carry out research following my interests. *Thank you very much for your academic and personal support!*

I am also grateful to Helmut Dietl for his interest in my work and for agreeing to be a member of the dissertation committee. Some of my articles have already greatly benefited from his precious feedback at seminars and conferences, and I am sure that they will further profit from his helpful critique and knowledgeable comments.

My thanks also goes to my former and current colleagues Marc Brechot, Rosa Cacabelos, Oliver Merz, Stephan Nüesch, Jan Pieper, Debora Rickenbacher, and Max Rüdisser, who have supported me during these years with friendship, encouragement, and inspiration. I also thank Fabio Banfi, Philippe Meier, and Michelle von Gunten for their excellent research assistance. Thanks also goes out to my colleagues from other chairs at the University of Zurich.

I am especially thankful to my colleague and co-author, Raphael Flepp, for providing an intense and fruitful collaboration. I greatly benefited from his rigorous approach to research and precise feedback at all times.

The first paper in this thesis has also benefited from the constructive review process at the *Journal of Economic Behavior & Organization*, for which I would like to thank the editor, William Neilson, and the anonymous referees. I am grateful to the All England Lawn Tennis Club, the French Tennis Federation, IBM, and the *Schweizer*

Radio und Fernsehen for providing the data I used in this dissertation. I am also indebted to Nico Gurtner from Mediapulse, Alex Frankel, Natalie Reid, and Mark Walker for useful consultation.

Last, I would like to thank my loved ones. I would like to express my most sincere gratitude to my parents Flavio and Rosanna and my sister Lisa for always believing in me, to my partner in life Laure for her unconditional support, and to all my friends for sharing with me all the beautiful things.

I take responsibility for any remaining errors.

Zurich, October 2017 Paolo Bizzozero

Framework Paper

As the title suggests, the aim of this dissertation is to test a number of economic theories by analyzing human behavior observed in sports-related settings. More specifically, the three articles in this dissertation provide important insights into the behavior of three different groups of economic agents—television viewers, professional sports players, and betting traders—using tennis as a research laboratory.

This dissertation contributes to a growing body of literature in the field of sports economics. Compared to other research fields like labor and health economics, sports economics is a relatively newer field within the broader discipline of economics (Santos & García, 2011). The first study using sports data that was published in a well-known economics journal appeared in the mid-1950s, when Rottenberg (1956) wrote an article about the characteristics of the labor market for baseball players. Some years went by before Neale (1964) published an influential article describing the peculiar economics of professional sports in another recognized journal.

In recent decades, the field of sports economics has significantly expanded its research agenda and has become more widely recognized. For example, the works of Wooders (2010), Pope and Schweitzer (2011), and Ertug and Castellucci (2013) have been published in leading scientific journals in economics and management sciences. Even if the reasons behind this expansion are manifold, sports economists recurrently mention two. First, the rapid growth in the demand for live sports events, for sports broadcasts, and for sporting goods and activities—which began in the 1970s in North America and in the 1980s in Europe—has resulted in an increased economic importance of the sports industry in modern society. Concurrent with this development, the interest of researchers and the research opportunities in this subject have rapidly expanded. By applying a "dose of economic thinking to the business of sports" (Fort, 2006), economists are actively contributing to a better understanding of the sports labor market, the competition for media rights, sporting governance, the demand and supply for sports products and services, and many other topics.

Second, researchers have realized that the scientific analysis of sports data can also contribute to the disciplines of economics, management science, and finance (Rosen & Sanderson, 2001). For example, economists view spectator sports as an ideal laboratory for empirical investigations because the outcomes are easily observable and measurable, the subjects are experienced professional players who have clear goals and incentives, the rules of the game are clearly defined, and the data are good and abundant (Palacios-Huerta, 2014). Furthermore, sports-related industries like betting, video gaming, and broadcasting have also grown, opening up interesting new areas of interdisciplinary research.

In line with the second perspective, the broad objective of this dissertation is to contribute to a better understanding of economics by analyzing sports and sports-related data using economic methods. This dissertation analyzes detailed tennis data combined with betting and TV audience data. Even though the sport of tennis is the common thread in all of the three articles presented below, it is not the focal point of this dissertation; rather, tennis serves as a research laboratory for studying human behavior in a real setting. Therefore, the reader of this dissertation does not need any specific knowledge of tennis to understand its content—specific tennis terminology will be explained when employed. Overall, our data give us the rare opportunity to explore a number of carefully selected economic questions, which are individually addressed in three empirical papers.

The first paper, *The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon* (see Appendix A.1—a version of this article has been published in the *Journal of Economic Behavior & Organization*), investigates the behavior of TV spectators while watching live tennis and asks how the amount of suspense and surprise at each moment of the match affect the entertainment value of the match.

Even though we frequently face or witness situations that are suspenseful (waiting for an important email) or surprising (receiving an unexpected gift), the empirical and formal economic literature on the topic is, at best, scarce. However, Ely, Frankel, and Kamenica (2015) recently contributed to filling the gap in the formal literature by introducing a theoretical framework in which a Bayesian audience derives entertainment utility (enjoyment) from *anticipated* changes in beliefs (suspense) and *actual* changes in beliefs (surprise).[1]

Our article contributes to filling the gap in the empirical literature by drawing on the model of Ely et al. and providing a novel setting in which we can measure the amount of suspense and surprise and link these measures to live TV audience figures. Tennis is an ideal laboratory because the level of suspense and surprise can drastically fluctuate throughout the match. For example, important points, comebacks, and spectacular rallies can make any moment more entertaining to watch. As viewers can easily and at no cost switch channels to maximize their utility from viewing, short-term variations in TV audience figures should reflect whether the audience is enjoying a given match (Alavy, Gaskell, Leach, & Szymanski, 2010).

Our tennis-related data facilitate our empirical analysis in two ways. First, we can operationalize suspense and surprise by using the in-play (during the match) betting odds originating from Betfair, one of the largest online betting exchanges worldwide. From the odds, we can derive the aggregate audience's belief at each point in the match about a given player's probability of winning the match by computing the inverse of the odds: for example, odds of 2.0 imply a 50% probability of winning the match. In our setting, higher suspense is attributed to greater variance in the

[1] In a Bayesian framework, probabilities quantify personal beliefs: people form hypotheses about the occurrence of specific events (e.g., "it will rain tomorrow") and attach probabilities to them based on their subjective levels of belief in these hypotheses ("with a 90% probability").

next period's belief, whereas greater surprise results from the occurrence of an event that strongly contradicts the audience's belief. Second, we can measure the demand for entertainment using live TV audience figures (ratings), which we match with the suspense and surprise measures.

We use over 8,500 minute-by-minute observations from 80 men's singles professional matches played at the Wimbledon Championships between 2009 and 2014. Our high-frequency panel data allow us to control for time-invariant factors that might jointly affect the audience level—like the stage of the competition, the day of the week, or the quality of the players—by using within-match variation. Importantly, our match fixed effects estimates show that both suspense and surprise are drivers of media entertainment demand, confirming the general intuition that both suspense and surprise are enjoyable features of entertainment content. Subsequent regression analyses reveal that surprise is more important in this regard than suspense and that both factors generate larger audiences during the later moments of a match. Finally, we show that these results are equally valid for male and female viewers.

Overall, our results are the first that provide empirical evidence that suspense and surprise drive media entertainment demand. Furthermore, our results have two important implications for the entertainment industry. First, entertainment content could be rigorously designed to increase suspense and surprise. Possible applications include, but are not limited to, movies, TV shows and series, and video games. The challenge remains in measuring people's beliefs in such contexts. Nonetheless, even for settings in which prediction markets are unavailable or where theoretical modeling is problematic, the measurement of beliefs is feasible. For example, one might analyze users' posts on social media platforms, such as Facebook or Twitter, using big data text analysis tools. Second, the timing and selling of TV advertisements could be improved. As suspense and surprise enhance entertainment and simultaneously attract human attention, commercials will be more effective following very suspenseful or surprising moments, thereby increasing the advertisement's effectiveness. Therefore, as an example, the price of a commercial during a TV broadcast like the Superbowl (the final game of the American football season, which in 2015 had 114 million U.S. viewers) could include a premium if shown just following a surprising moment like a touchdown.

Finally, there are plenty of open research questions to address in this subject. For example, using individual-level data, one could investigate how different types of consumers react to suspense and surprise; or, using longer time-series data, one could investigate how specific time patterns of suspense and surprise affect media consumption.

The second paper, *When do professionals play minimax? Evidence from the tennis court* (see Appendix A.2), analyzes the behavior of professional tennis players when playing at one of the most prestigious tennis tournaments worldwide, the Wimbledon Championships, and asks whether and when (under what conditions) professionals play as predicted using John von Neumann's minimax, a fundamental game theory theorem.

In a seminal paper, O'Neill (1987) stresses the importance of testing the empirical validity of the minimax theorem because many economic models are based on it or its generalization, the Nash equilibrium. In general, in experimental studies, students who played simple games such as matching pennies or card games do not play as predicted by the minimax: they do not play according to the equilibrium mixtures and their choices are not serially independent (e.g., Rapoport & Boebel, 1992; Binmore, Swierzbinski, & Proulx, 2001; Camerer, 2003; Levitt, List, & Reiley, 2010). In contrast, previous investigations have established that the strategic behavior of players who are both experienced and have high incentives (like professional sport players) is somewhat closer to that predicted by the minimax theorem (Palacios-Huerta & Volij, 2008).

Tennis data provide a valuable opportunity for empirically testing the minimax theorem. We use professional tennis and model the tennis serve as a two-person zero-sum simultaneous mixed-strategy game: in this game, the server aims to maximize the expected probability of scoring a point on the serve, and the receiver aims to minimize this probability. Both the server and the receiver have two strategies: to play left or right. Even though this model simplifies the available strategies that both players have—like the serve speed, depth, and type (slice, topspin, etc.)—it provides an important opportunity to test the two minimax predictions in a real setting.

Walker and Wooders (2001) analyze 10 men's matches and find that the players' behavior is consistent only with prediction (1), i.e., that the winning probabilities across choices are equal, whereas Hsu, Huang, and Tang (2007) analyze 10 different men's matches and show that the players' behavior is consistent with prediction (1) and prediction (2), i.e., that players' choices are random. In contrast to these studies, we use a larger and significantly more heterogeneous sample of tennis matches and show that the players' behavior is consistent only with prediction (2).

A common property of Walker and Wooders (2001) and Hsu et al. (2007) is that these studies both focus only on testing whether the observed behavior of professional athletes is "close enough" to the theoretical predictions to be considered consistent with minimax. However, they do not address the important question as to when—under which conditions—players deviate more (or less) from the minimax predictions. Such an investigation is important for understanding how the players' behavior differs under, for example, stress or fatigue, or when playing against opponents with different skills.

This article contributes to filling this gap by showing that professionals have a tendency to play toward minimax in matches between right-handed players (versus matches with a left-handed player), in the later sets of a match, in shorter matches, and especially in unbalanced matches (between players with different skills). Most importantly, further analyses show that the behavior of top-ranked servers is less consistent with minimax when they face non-top-ranked receivers. While this result appears surprising at first glance, we can show that top servers would decrease their likelihood of winning a match by playing closer to minimax against non-top receivers; therefore, deviating is a rational response.

Taken together, our article makes three contributions. First, we provide a more powerful test of the predictions of the minimax theorem using a significantly larger and more heterogeneous sample than previous studies. Second, we show under which specific conditions the players' behavior is closer to minimax. Here, further research is needed to clarify how the specific conditions of tennis are generalizable: for example, how do our results for the "later sets of a match" subsample apply to general situations under pressure? Third, our results underline the importance of controlling for differences in the players' skill levels. Therefore, future lab experiments should determine whether the participants have similar or different skill levels for the games that are played.

The third paper, *What is the effect of insider trading on price efficiency? Evidence from a betting exchange* (see Appendix A.3), studies the behavior of a group of betting traders who have a fleeting informational advantage with respect to other traders, so-called insider traders. This paper asks to what extent insider trading affects the efficiency of security prices on a betting exchange.

Researchers and regulators have long debated about both the fairness and the economic implications of insider trading. Concerning the latter, which is the focus of our article, the key issue involves the assessment of the impact of insider trading on price efficiency. However, little empirical evidence has been collected because insider trading on material, pricing-relevant information is illegal in most financial markets. The only exceptions are the works of Cornell and Sirri (1992), Meulbroek (1992), and Chakravarty and McConnell (1997), which analyze the market's reaction to illegal insider trades (the data originate from criminal and civil litigation reports) and show that insider trades lead to more rapid price discovery. However, Chakravarty and McConnell (1999) disputed these studies by showing, using a refined methodology to re-analyze the data of those previous studies, that the effect of insider trades on prices is not discernibly different from that of non-insider trades. Thus, the evidence on the topic is generally limited and mixed. Our paper's contribution is that it provides a simpler setting that allows us to overcome the difficulties of previous investigations.

One key advantage of our in-play tennis-betting setting comes from the inevitable technical delay in the transmission of match information from the stadium to end receivers. Media studies, bettors, and Betfair indicate that the "live" TV images and internet scoreboards are delayed by *at least* five seconds. Insider traders sitting in the stadium, so-called courtsiders, have an informational advantage as they observe important information at least five seconds before outsiders. Therefore, any price movement observed within five seconds after an important informational event can be attributed to insider trading activity.

To quantify the impact of insider trading activity on the price discovery after important news events, we use detailed second-by-second data from two major tennis tournaments, the French Open and the Wimbledon Championships, over the 2009–2014 period. An important event in tennis is the end of a set because winning a set constitutes a decisive step toward victory. Knowing which player wins the set earlier than the public means holding an informational advantage since that information is material.

Using event study methods, we find that the cumulative abnormal returns averaged across the 365 event observations start increasing immediately after the set event time, which demonstrates the existence of insider traders, and stabilizes after approximately seven seconds. Most importantly, we show that the cumulative abnormal return during the first five insider trading seconds following the set events—when the TV and other slower traders have not yet seen the event—is more than 60% of the full price reaction observed once the public receives the new information. Furthermore, we show that the the impact of insider trading is even larger for unanticipated news events, like tie-break set events, when inside information is more valuable. Finally, we estimate that a simple dynamic back-lay trading strategy implemented in the seconds after the event yields large risk-free profits to insiders, varying between 5% and 7%.

Overall, our results are important because they provide empirical evidence from a simple but advantageous setting that insider trading significantly contributes to higher price efficiency and thus market quality. Insider traders have large financial incentives to rapidly integrate new information into the market. Overall, when debating the effects of insider trading on a betting exchange as well as on any financial market, the negative externalities from the adverse selection costs to slower traders should be weighed against the positive externalities from greater price efficiency.

In conclusion, this dissertation shows—in addition to the individual contributions of the three papers—how the economic analysis of sports and sports-related data can contribute novel scientific knowledge. Whenever appropriate data in classical economic, financial, or managerial areas are difficult to find or are unsuitable, researchers should look at the world of sports because of the availability of abundant high-quality data. The following three articles contribute to the emerging strand of literature on sports economics whose goal is to broaden our understanding of economics through the analysis of human behavior in sports and sports-related areas. The papers also put forwards a number of opportunities for further research in the hope that more researchers will join this exciting field.

References

Alavy, K., Gaskell, A., Leach, S., & Szymanski, S. (2010). On the edge of your seat: Demand for football on television and the uncertainty of outcome hypothesis. *International Journal of Sport Finance*, *5*(2), 75–95.

Binmore, K., Swierzbinski, J., & Proulx, C. (2001). Does minimax work? An experimental study. *The Economic Journal*, *111*(473), 445–464.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction.* Princeton University Press.

Chakravarty, S., & McConnell, J. J. (1997). An analysis of prices, bid/ask spreads, and bid and ask depths surrounding Ivan Boesky's illegal trading in Carnation's stock. *Financial Management*, 18–34.

Chakravarty, S., & McConnell, J. J. (1999). Does insider trading really move stock prices? *Journal of Financial and Quantitative Analysis*, *34*(2), 191–209.

Cornell, B., & Sirri, E. R. (1992). The reaction of investors and stock prices to insider trading. *The Journal of Finance*, *47*(3), 1031–1059.

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, *123*(1), 215–260.

Ertug, G., & Castellucci, F. (2013). Getting what you need: How reputation and status affect team performance, hiring, and salaries in the NBA. *Academy of Management Journal*, *56*(2), 407–431.

Fort, R. D. (2006). *Sports economics.* Prentice Hall, Upper Saddle River, New Jersey.

Hsu, S.-H., Huang, C.-Y., & Tang, C.-T. (2007). Minimax play at Wimbledon: Comment. *The American Economic Review*, *97*(1), 517–523.

Levitt, S. D., List, J. A., & Reiley, D. H. (2010). What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments. *Econometrica*, *78*(4), 1413–1434.

Meulbroek, L. K. (1992). An empirical analysis of illegal insider trading. *The Journal of Finance*, *47*(5), 1661–1699.

Neale, W. C. (1964). The peculiar economics of professional sports. *The Quarterly Journal of Economics*, *78*(1), 1–14.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences*, *84*(7), 2106–2109.

Palacios-Huerta, I. (2014). *Beautiful game theory: How soccer can help economics.* Princeton University Press.

Palacios-Huerta, I., & Volij, O. (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica*, *76*(1), 71–115.

Pope, D. G., & Schweitzer, M. E. (2011). Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *The American Economic Review*, *101*(1), 129–157.

Rapoport, A., & Boebel, R. B. (1992). Mixed strategies in strictly competitive games: A further test of the minimax hypothesis. *Games and Economic Behavior*, *4*(2), 261–283.

Rosen, S., & Sanderson, A. (2001). Labour markets in professional sports. *The Economic Journal*, *111*(469), 47–68.

Rottenberg, S. (1956). The baseball players' labor market. *Journal of Political Economy*, *64*(3), 242–258.

Santos, J. M. S., & García, P. C. (2011). A bibliometric analysis of sports economics research. *International Journal of Sport Finance*, *6*(3), 222.

Walker, M., & Wooders, J. (2001). Minimax play at Wimbledon. *The American Economic Review*, *91*(5), 1521–1538.

Wooders, J. (2010). Does experience teach? Professionals and minimax play in the lab. *Econometrica*, *78*(3), 1143–1154.

Appendix: Papers included in this dissertation

A.1

The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon[†]

Paolo Bizzozero, Raphael Flepp, Egon Franck

Abstract

This paper empirically examines how suspense and surprise affect the demand for entertainment. We use a tennis tournament, the Wimbledon Championships, as a natural laboratory. This setting allows us to both operationalize suspense and surprise by using the audience's beliefs regarding the outcome of the match and observe the demand for live entertainment using TV audience figures. Our match fixed effects estimates of 8,563 minute-by-minute observations from 80 men's singles matches between 2009 and 2014 show that both suspense and surprise are drivers of media entertainment demand. In general, surprise seems to be more important in this regard than suspense, and both factors matter more during a match's later moments. We discuss important implications for the design of entertainment content to maximize entertainment demand.

JEL Classification: D83, L82, L83

Keywords: Suspense, surprise, entertainment, TV audience, betting odds, tennis

[†]Reprinted from *Journal of Economic Behavior & Organization*, *130*, Bizzozero, P., Flepp, R., & Franck, E., The importance of suspense and surprise in entertainment demand: Evidence from Wimbledon, 47–63, Copyright (2016), with permission from Elsevier. 10.1016/j.jebo.2016.07.006

1 Introduction

Media entertainment plays an important role in people's daily lives. Vorderer, Klimmt, and Ritterfeld (2004) describe media entertainment as enjoyment from consuming media content, whether at home or at an outside venue. Given that entertainment providers are facing stiffer competition in the entertainment market, understanding precisely what factors drive the demand for entertainment content is of critical importance.

Previous studies have identified *suspense* and *surprise* as two major determinants of enjoyment associated with media consumption (e.g., Zillmann, 1991, 1996; Vorderer et al., 2004). The online Cambridge English Dictionary defines suspense as "a feeling of excitement or anxiety while waiting for something uncertain to happen" and surprise as "an unexpected event, or the feeling caused when something unexpected happens." Importantly, both occur exclusively in situations in which there is concern over uncertain outcomes (Comisky & Bryant, 1982).

Suspense and surprise are best understood and modeled in a Bayesian setting (Ely et al., 2015). In this setting, probabilities quantify personal beliefs: people form hypotheses about the occurrence of specific events (e.g., "it will rain tomorrow") and attach probabilities to them based on their subjective levels of belief in these hypotheses ("with a 90% probability"). In the Bayesian view, people will transform their *prior* beliefs into *posterior* beliefs when new and relevant information arrives (Itti & Baldi, 2009). This continuous process of forming and updating beliefs leads to entertainment based on the experience of suspense and surprise, where suspense and surprise are the forward- and backward-looking emotions, respectively.

Suspense evolves through the assessment of future events, with a moment carrying more suspense when some crucial uncertainty is soon to be resolved (Vorderer et al., 2013), such as a researcher opening a letter with the committee's decision on his or her research grant application. By contrast, surprise evolves by assessing past events, with a moment carrying more surprise immediately after an unexpected event occurs (Itti & Baldi, 2009), such as after an underdog soccer team scoring the winning goal.

Although it is intuitive that suspense and surprise matter in the context of entertainment, empirical tests are difficult to design because people's beliefs and their enjoyment are hard to observe. Moreover, little is known about the importance of suspense relative to surprise or about their importance with respect to the passage of time. In this paper, we address these questions by employing high-frequency data from a tennis tournament, the Wimbledon Championships, which offers our research two unique advantages.

The first advantage is that we can quantify the audience's beliefs because modeling tennis situations is possible. In tennis, a Bayesian audience forms beliefs about the final outcome of the match, i.e., about the likelihood that a particular player will win a particular match.[1] We estimate the relevant beliefs at the point-by-point level in two ways: first, we use a Markov model that requires the player's probability of winning a service point and the current score as inputs; second, we use in-play betting odds.

The second advantage that tennis offers is that the demand for entertainment is observed using high-frequency minute-by-minute live TV audience figures (*ratings*) during the matches. As viewers can easily—and at no cost—switch channels or turn off the TV to maximize their utility from viewing, short-term variations in TV audience figures reflect whether the audience is enjoying a given match (Alavy, Gaskell, Leach, & Szymanski, 2010). By using minute-by-minute information regarding aggregate viewers' behavior, we can uncover an audience's underlying preferences for entertainment in a real-world environment.[2]

[1] The same idea can be applied in other settings. For example, people assign probabilities to the hypothesis that a president will be reelected, that a mission will succeed, or that a company's earnings will beat analysts' consensus estimates. What differentiates tennis from other settings is the frequency with which events happen and new information is revealed.

[2] TV remains the central provider of entertainment content despite the increasing supply of entertainment available on the Internet. According to the U.S. Bureau of Labor Statistics, individuals aged 15 and over watched TV for 2.8 hours per day on average in 2013, accounting for more than half of their leisure time.

This paper contributes to the literature by presenting an analysis of unique and naturally occurring field data that provide a rare opportunity to empirically investigate the importance of suspense and surprise when consuming a media entertainment product. Our empirical analyses reveal that both suspense and surprise have a positive effect on entertainment demand. Using 8,563 minute-by-minute observations from 80 men's singles matches between 2009 and 2014, our match fixed effects estimates reveal that minutes with more surprise and suspense have significantly higher live TV ratings. This result indicates that suspense and surprise are complementary and that demand for entertainment is stronger for higher levels of suspense and surprise. In particular, a one standard deviation increase in suspense (surprise) is associated with an audience increase of approximately 1,200 (2,200) viewers per minute. For some perspective, the minute-level effect of a one standard deviation increase in suspense and surprise combined corresponds roughly to a 3% audience increase (based on an average audience of approximately 100,000 viewers in our sample). Although we cannot compare our results with those from previous studies, our estimates suggest that the impact of suspense and surprise on TV audience figures is economically non-trivial.

Moreover, we find that the audience impact of surprise is consistently greater than that for suspense: depending on the model used for computing the audience's beliefs, the estimated effects for surprise are between two and five times greater than those for suspense. Hence, surprise appears to be more important than suspense in entertainment demand. In addition, over the course of a match, the impact of both suspense and surprise clearly increases. This implies that the entertainment effect of suspense and surprise is larger when the stakes are higher.

To the best of our knowledge, this paper is the first to test Bayesian theory on suspense jointly with surprise under natural conditions. We provide a framework that entertainment industry managers can use to measure an audience's beliefs, which can then be used to measure entertainment from suspense and surprise. Although in tennis there is not much room for artificially increasing suspense and surprise, the implications of our study are far more important for other entertainment settings in which content can be designed ad hoc to increase the public's enjoyment. Designers of films, TV series, TV shows, online videos, novels, or gambling games should be aware of people's preferences for suspense and surprise, their increasing significance towards the end of a media event, and the greater importance of surprise.

The remainder of this paper is organized as follows. Section 2 reviews the literature. Section 3 describes our setting and data. Section 4 outlines the operationalization of suspense and surprise and our empirical methodology. Section 5 presents the empirical results and various robustness checks. Section 6 discusses several implications and concludes.

2 Literature review

The theoretical literature on suspense and surprise is limited. Yet Ely et al. (2015) recently filled this gap by introducing a framework in which a Bayesian audience derives entertainment utility (enjoyment) from *anticipated* changes in beliefs (suspense) and *actual* changes in beliefs (surprise). In the model developed by these authors, higher suspense results from greater variance in the next period's beliefs—what is currently happening versus what is expected to happen next—and higher surprise results from greater distance between previous and current beliefs.

However, no study has yet empirically investigated the relationship between suspense, surprise and enjoyment. Most of the relevant studies are laboratory experiments that focus either on suspense or surprise. For example, Bryant et al. (1994) recorded a football match and manipulated its commentary to create a high-suspense version and a low-suspense version. After having watched one of these two versions, participants were given a questionnaire and asked to rate their enjoyment on a scale from 0 to 10. The results from Bryant et al. (1994) show that viewing the high-suspense version was significantly more enjoyable and exciting. Su-lin et al. (1997) and Peterson and Raney (2008), using students as participants, examine suspense as a factor in the enjoyment of National Collegiate Athletic Association (NCAA) men's basketball games. Both studies operationalize suspense as the final point differential in a game and enjoyment as the average of seven different items (e.g., "the game excited me" or "I enjoyed the game") that are rated on a scale from 0 to 10. Their results show that higher suspense leads to greater enjoyment.

In an experimental setting, Itti and Baldi (2009) test whether surprise attracts the attention of participants when watching one of several videoclips, including television broadcasts, such as news, sports, commercials, and outdoor scenes. They define surprise in Bayesian terms as the distance between the prior and posterior distributions of beliefs, and they employ eye-tracking technology to measure attention. Their results demonstrate that surprise explains the greatest portion of human eye movements, indicating that humans are attracted to surprising elements in video displays.[3] Moreover, in a laboratory setting, Alwitt (2002) finds that viewers perceive suspenseful TV commercials as shorter, attributing this effect to the viewers' intensified attention and interest.

In the sports economics literature, in particular, suspense has been often associated with "uncertainty of outcome." The uncertainty of outcome hypothesis posits that more uncertainty about the final winner of a sports competition leads to more suspense (Borland & MacDonald, 2003). Alavy et al. (2010) use minute-by-minute audience figures from 248 English Premier League matches to measure the effect of outcome uncertainty on entertainment demand. They find that matches with less probability of ending in a draw—but also with less score differential between teams—generate more viewers. In a sports-related article, Olson and Stone (2014) model viewers' entertainment as a function of suspense to evaluate whether the introduction

[3] Baldi and Itti (2010) build on these results and provide further evidence of this relationship. For a recent review of studies on the surprise-attention link, see Horstmann (2015).

of post-season playoffs in U.S. college football would be an improvement over the current system. Using match-level Nielsen ratings over 2011–2013 for 70 football and basketball matches, they show that the level of viewers' entertainment significantly increases with the championship's suspense level.

3 Setting and data

Clearly, a tennis match provides many moments of various levels of suspense and surprise: comebacks, break points, tie-breaks, injuries, and spectacular rallies can make any moment entertaining, whereas unimportant points can make any moment less entertaining to watch.[4] We use the 2009 Wimbledon men's final between Andy Roddick and Roger Federer as our illustrative example. After winning the first set, Roddick had four set points in the second set, putting him only one set away from the championship. However, supported by his strong service, Federer won all of Roddick's set points and eventually won the set. Because the audience had to strongly readjust its beliefs about Federer's chances of winning, we describe such circumstances as surprising. In the final set, which is played until one player wins at least six games by at least a two-game spread, Federer finally won 16–14 in a set that lasted more than 90 minutes. Because each point could potentially bring a player's winning probability very close to either zero or one, we describe such circumstances as suspenseful.

Our definition of suspense and surprise is best understood in a Bayesian framework. Simply put, a Bayesian audience has some current beliefs, based on the information currently available, about a specific uncertain outcome. Upon the arrival of new, relevant information, the audience will update its beliefs, which are then called posterior beliefs. In tennis, viewers form and continuously update their beliefs about the "hypothesis" that a given player might win the match. To quantify moments of various

[4] In the Appendix A.1, Part I briefly describes the rules of tennis and its jargon.

levels of suspense and surprise, we must therefore estimate the relevant beliefs at the point-by-point level. We do this using two methods.

In the Markov chain method, we rely on the explicit structure of the data-generating process in tennis. In tennis, points are linked to games, games to sets, and sets to matches; thus, a match can be modeled as a binary Markov chain (Newton & Aslam, 2009; O'Malley, 2008).[5] We estimate the unique belief path for each match using a computer program that computes the likelihood of winning the match for a given player point-by-point over the match.[6] The only input for this simple model is the score and the probability of winning a point on serve. Detailed match data at the point level are provided by IBM, the official supplier of information technology to the Wimbledon Championships. Beyond general information about the match, such as the players, courts, start and end match times, these data also contain point-by-point information on the current score, time (exact to the second), server, and winner.

In the betting odds method, we rely on the information content of in-play betting odds from *in-play* betting markets. Wolfers and Zitzewitz (2006) show that betting odds provide valuable estimates of average, aggregate beliefs about the probability that an event will occur. The odds originate from Betfair, one of the largest online betting markets, and are provided by Fracsoft, a data vendor.[7] Betfair's online platform provides a market for opinions and for participants to bet against one another by offering and accepting odds under which they are willing to buy or sell a certain bet. Bettors mostly follow the match live on TV or on other electronic devices and continuously place their bets during the match: whenever new information becomes observable, they update their beliefs, and the odds change accordingly.

[5] Walker et al. (2011, p. 490) illustrate the binary Markov scoring rule for a game of tennis. Moreover, Liu (2001) and Barnett and Clarke (2005) propose similar ways of modeling the probability of winning a match.

[6] The software program, described in Klaassen and Magnus (2014), is called "Richard" and is freely available online at http://www.janmagnus.nl/misc/wimbledon.pdf with detailed instructions.

[7] Trading volumes on Betfair are very large. For example, 1.2 billion bets were placed in 2014, resulting in a total trading value of roughly $92 billion. Croxson and Reade (2014) estimate that the daily trading intensity on Betfair during the 2005–2007 period was greater than the daily trading intensity on all the major European Stock Exchanges combined. With regard to our sample, we observe an average total trading volume of $28.5 million per match, 70% of which was placed in-play. Approximately $78 million worth of bets were placed on the 2014 Wimbledon men's final alone, and 92% of that amount was placed in-play.

Importantly, as the Markov belief is based on the actual score and on the server's probability of winning a point, the in-play betting odds belief should be more accurate. In fact, odds would reflect not only the newest information, such as a player's injury, but also a number of historical factors, such as a player's performance record on grass courts or previous head-to-head records between the players. In this sense, we consider the betting odds method as the main specification.[8] The odds enable us to derive the aggregate market's belief at each point in the match about a given player's probability of winning the match. Indeed, the inverse of the odds on the expected match winner can be interpreted as the aggregate current belief about a player's match winning probability (Hasbrouck, 1991).

To measure entertainment demand, we gather high-frequency TV audience ratings over the 2009–2014 period on all Wimbledon men's singles matches that were transmitted live on the Swiss national German-language channels *Schweizer Fernsehen Zwei* (SRF2) and *Schweizer Fernsehen Info* (SRFinfo), two of the largest Swiss broadcaster's free channels.[9] In Switzerland, Wimbledon—and tennis as a whole—enjoys good TV coverage. Overall, SRF broadcasted 108 Wimbledon men's singles matches between 2009 and 2014. In our analysis, we exclude 28 matches: three matches are excluded because no betting data is available, whereas 25 partially and shortly transmitted matches are excluded because of potential spillover effects in the audience ratings caused by the preceding and following TV programs.[10]

Mediapulse, a Swiss ratings firm essentially equivalent to Nielsen, generates audience statistics using survey data from a panel encompassing 1,870 households across

[8] We thank an anonymous referee for this suggestion.

[9] Whereas SRF2 focuses on either live or recorded sports programming, SRFinfo chiefly rebroadcasts programs from SRF1 (the first national channel) and SRF2. However, it also occasionally acts as a complementary channel for live sportscasts in the event of programming conflicts. On average, in our sample, SRF2 has $112,515$ viewers, whereas SRFinfo has $45,853$ viewers.

[10] Although the average duration of the excluded matches is only 14 minutes, one might worry that the exclusion of 25 matches introduces a selection bias because the broadcaster might stop broadcasting matches with low suspense, low surprise, or both. As we cannot compute the levels of suspense and surprise due to IBM data unavailability, we examine the TV channel content description (the list is available upon request) and also directly asked the broadcaster. We discovered that the broadcaster cannot observe live audience figures: as a result, programming decisions do not depend on such live audience figures. It rather appears that the broadcaster filled programming "gaps" with some scenes from these 25 matches.

Switzerland, which contains approximately 4,200 people three years of age and over. As the advertising industry also uses Mediapulse's data, the panel must meet strict requirements to reflect the Swiss population as accurately and representatively as possible (see Appendix A.2 for further details). TV audience ratings measure the total number of single viewers watching the channel at each moment. For example, a rating of 122,500 indicates that 122,500 single viewers were tuned into the program on average during a particular minute. For any minute when more people tune into the tennis match (either from another channel or by turning on the TV) than turn off the match, the rating will increase.

Our final data set consists of roughly 8,500 minute-by-minute observations on detailed match statistics, in-play betting odds, and live TV ratings. In Table 1, Panel A reports descriptive statistics for the dependent variable *audience* on the 80 matches in the final sample. An average match has an audience of slightly more than 106,000 spectators, corresponding to a market share of 16.9%. Panel B additionally shows that *audience* is larger at later tournament stages. Our sample is heterogeneous, containing matches from the first stage up to the finals and a total of 58 unique players. An average match lasts 150 minutes, consists of 220 points and 3.6 sets, and each set consists of 10 games. Late tournament matches are longer, as they typically are more balanced.

4 Estimation approach

4.1 Suspense and surprise

The Markov method

The construction of suspense and surprise closely follows the work of Ely et al. (2015), in which the entertainment utility of the Bayesian audience is a function of the Markov belief path. First, we model suspense in the form of an expectation, where higher

Table 1
Descriptive statistics of the audience and sample characteristics.

Panel A: Descriptive statistics						
Variable	Description	N	Mean	Std. dev.	Min	Max
audience	TV Rating	8,563	106.47	136.81	1.77	1,083.7
Panel B: Sample characteristics						
Tournament stage	N	Length (minutes)	Points in match	Sets in match	Games in set	*audience*
$1st$ stage	12	117	186	3.2	9.4	44.0
$2nd$ stage	14	121	189	3.2	9.6	55.2
$3rd$ stage	12	147	223	3.7	9.4	75.8
$4th$ stage	10	155	220	3.6	9.9	83.2
Quarterfinal	14	167	240	3.8	10.2	87.4
Semifinal	12	169	236	3.7	10.5	89.1
Final	6	202	274	4.0	11.3	362.8
Full sample	80	150	220	3.6	10.0	106.4

Notes: The table reports descriptive statistics for the 80 men's singles matches played at Wimbledon between 2009 and 2014. Panel A describes the TV audience ratings (in thousands, except for N), whereas Panel B provides the means of additional match characteristics (by tournament stage and for the full sample).

suspense is attributed to greater variance in the next period's belief.[11] For the point p:

$$SUS_p^{Markov} = [\mathrm{E}\sum_{\omega}(\mu_{p+1}^{\omega} - \mu_p)^2]^{1/2}, \tag{1}$$

where μ_p refers to the current player's probability of winning the match (the current belief) at the moment when point p is scored and μ_{p+1}^{ω} refers to the anticipated posterior probability of a player's winning the match. The posterior belief depends on the realization of state ω, which in this setting is a binary variable: $\omega = 1$ when the server i wins the point (with probability S_i) or $\omega = 0$ when he loses it (with probability $1 - S_i$). Thus, we can rewrite Eq. (1) as follows:

$$SUS_p^{Markov} = [S_i \cdot (\mathrm{E}[\mu_{p+1}^1] - \mu_p)^2 + (1 - S_i) \cdot (\mathrm{E}[\mu_{p+1}^0] - \mu_p)^2]^{1/2}. \tag{2}$$

Because only one state occurs in reality, we also estimate the *counterfactual* posterior belief, defined as the probability of winning the match for the unobserved state. To do so, we replace the actual score with the counterfactual score in the Markov model. For example, if player i actually serves and wins the first point of the match ($\omega = 1$),

[11]In the Appendix A.3, Part I illustrates how we compute SUS_p^{Markov} and SUR_p^{Markov} with a numerical example.

the counterfactual belief is computed by assuming that he *lost* the first point.[12] The player's probability of winning a service point (S_i) is computed for each player in our sample based on historical data from the Association of Tennis Professionals (ATP) website.[13] To obtain accurate predictions from the model, we set μ_0, i.e., the winning probability at the beginning of the match ($p = 0$), equal to the corresponding odds-implied probability.[14]

Second, we model surprise in the form of the Euclidean distance between the prior belief and the current belief, where greater surprise results from the occurrence of an event that strongly contradicts the audience's belief, constraining the audience to change its beliefs (Itti & Baldi, 2009). For the point p:

$$SUR_p^{Markov} = |\mu_p - \mu_{p-1}|, \tag{3}$$

where μ_{p-1} refers to the probability of winning the match for a player one point earlier.

The betting odds method

The construction of suspense and surprise based on betting odds is straightforward and similar to the Markov method. Following Hasbrouck (1991), we compute the average mid-odds from the best buy $odds_{ip}^{back}$ and sell $odds_{ip}^{lay}$ for player i for each point p in match m as follows:

$$odds_{ip}^{mid} = \frac{odds_{ip}^{back} + odds_{ip}^{lay}}{2} \quad \forall(p, m), \tag{4}$$

from which we derive the implied winning probability for player i:

[12]According to Eq. (2), if the server actually wins (loses) the next service point, the counterfactual is represented by $\mathrm{E}[\mu_{p+1}^0]$ ($\mathrm{E}[\mu_{p+1}^1]$).

[13]In the Appendix A.1, Part II provides further details regarding how we compute S_i.

[14]We are grateful to an anonymous referee for this suggestion.

$$\tilde{\nu}_{ip} = \frac{1}{odds_{ip}^{mid}} \quad \forall(p, m). \tag{5}$$

Although the sum of the winning market probability for player i and player j ($\tilde{\nu}_{ip}+\tilde{\nu}_{jp}$) should sum up to one in a frictionless market, in practice it rarely does so because of transaction costs. Following the standard approach to eliminating this overround (e.g., Forrest et al., 2005), we adjust the implied winning probability to obtain the final market implied winning probability for player i as follows:

$$\nu_{ip} = \frac{\tilde{\nu}_{ip}}{\tilde{\nu}_{ip} + \tilde{\nu}_{jp}} \quad \forall(p, m). \tag{6}$$

As the change in the implied winning probability is symmetric, the choice of which player's probability of winning to consider is irrelevant, so we can drop the player subscript i.

First, we define surprise for point p as follows:

$$SUR_p^{odds} = |\nu_p - \nu_{p-1}|, \tag{7}$$

where ν_p refers to the current player's probability of winning the match (the odds-implied current belief) at point p and ν_{p-1} refers to the odds-implied prior belief at the point $p-1$. The only difference from Eq. (3) involves the type of data used.

Second, as we cannot implement the baseline metric of suspense by relying entirely on betting odds, we redefine the baseline suspense measure, but only slightly. By definition, betting odds reflect bettors' *current* beliefs at any point in the match, based on a given information set. However, we do not know which value the odds would have taken had the next point gone differently, i.e., the counterfactual posterior belief. Therefore, we come up with a "hybrid" implementation of the baseline suspense measure in which we use the Markov chain model only to determine the counterfactual posterior belief ($\mathrm{E}[\mu^{\omega}_{p+1}]$) for the unobserved state. Because the audience must also estimate the counterfactual probability, this procedure is suitable for our purposes. Analogously to Eq. (2), we model suspense for point p as follows:

$$SUS^{odds}_p = \begin{cases} [S_i \cdot (\nu_{p+1} - \nu_p)^2 + (1 - S_i) \cdot (\mathrm{E}[\mu^0_{p+1}] - \nu_p)^2]^{1/2} & \text{if } \omega = 1, \\ [S_i \cdot (\mathrm{E}[\mu^1_{p+1}] - \nu_p)^2 + (1 - S_i) \cdot (\nu_{p+1} - \nu_p)^2]^{1/2} & \text{if } \omega = 0. \end{cases} \quad (8)$$

Against the backdrop of Eq. (2), we now substitute the Markov current belief μ_p with the odds-implied current belief ν_p. Concerning the posterior beliefs, we must distinguish between the states ω: when the server wins the service point ($\omega = 1$), we substitute $\mathrm{E}[\mu^1_{p+1}]$ with the actual odds-implied posterior belief ν_{p+1}, and we use the Markov posterior belief for the counterfactual ($\mathrm{E}[\mu^0_{p+1}]$); when the server loses the service point ($\omega = 0$), we substitute $\mathrm{E}[\mu^0_{p+1}]$ with the actual odds-implied posterior belief ν_{p+1}, while we use the Markov posterior belief for the counterfactual ($\mathrm{E}[\mu^1_{p+1}]$).

Finally, because we observe minute-by-minute variation in the TV audience, we translate suspense and surprise from point level to minute level by computing the average suspense or surprise over the minute during which more than one point is scored within a minute.[15]

A comparison of the Markov and the betting odds beliefs

Fig. 1 shows the belief paths regarding Novak Djokovic's chance to win the match he played against Roger Federer on 7th June, 2014. The black line shows the belief path computed using the betting odds, whereas the grey line shows the Markov belief path. Both beliefs start at 60%, the implied probability of winning from the in-play betting odds—apparently, the bettors thought Djokovic was the favorite—and end at 100% for Djokovic, who won the final in five sets.

Notably, the belief path from the Markov model highly correlates (0.991) with the be-

[15]We also test an alternative method that considers only the last point scored in any minute when more than one point is scored. This alternative does not alter our results.

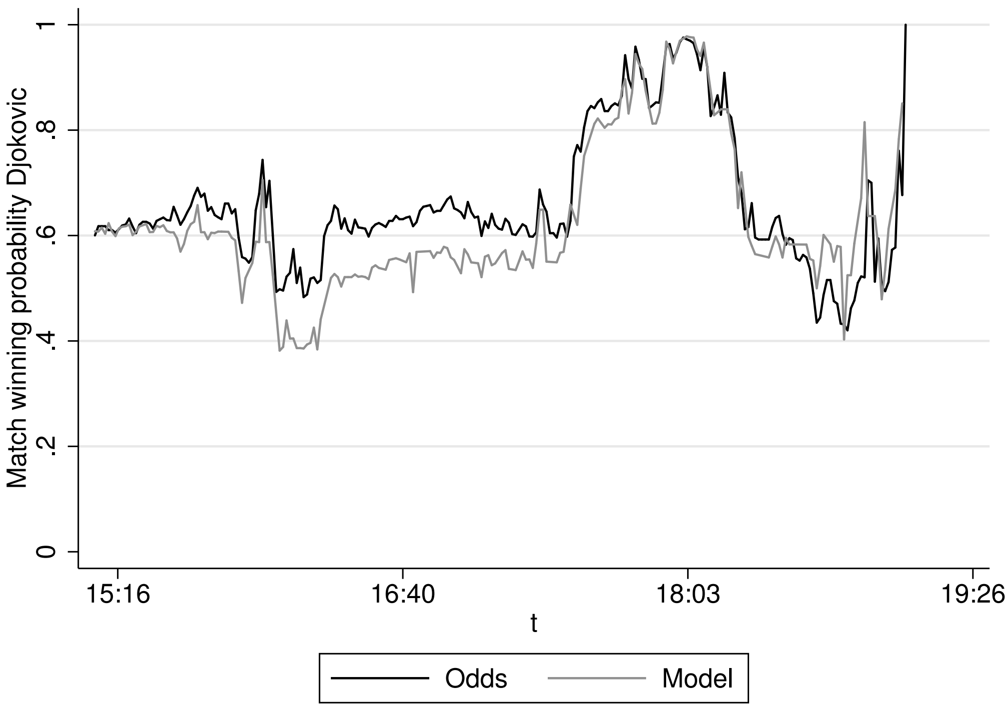


Figure 1
Displayed are the belief paths about Novak Djokovic's chance to win his 7th June, 2014 match against Roger Federer. Djokovic won in five sets with the score: 6–7; 6–4; 7–6; 5–7; 6–4. A total of $72 million was bet *in-play* on this Wimbledon final.

lief path extracted from the betting odds. Allegedly, professional bettors also strongly rely on computer programs to estimate the underlying player's winning probability during a match (Hutchins, 2014). Nonetheless, the Markov and betting odds belief paths might slightly diverge due to the arrival of new, relevant information beyond server and score, such as the signs of an injury, weather changes, or the stadium atmosphere.[16] However, because suspense and surprise are computed from *changes* in beliefs, the size of the vertical gap between the odds belief and the Markov belief is irrelevant—it is much more important that the beliefs are positively correlated.

Table 2 reports summary statistics for the suspense and surprise variables. Means

[16] Another cause might rely on match-fixing. In January 2016, BBC and BuzzFeed News uncovered evidence of widespread suspected match-fixing in tennis, including some matches at Wimbledon (see, for example, the article of Simon Cox at www.bbc.co.uk/sport/tennis/35319202). Because match-fixing could bias the betting odds, it could also bias the odds-derived belief of a player's winning probability. However, we believe that our results are not systematically affected for two main reasons: first, the report refers to events from approximately ten years earlier (two matches listed in the report are from 27 June 2006 and 26 June 2007), i.e., a period not covered by our sample, and second, although some of our matches may have been the subject of match-fixing without our knowledge, our suspense and surprise measures are based on *changes* in the winning probability, not on absolute levels.

Table 2
Descriptive statistics of suspense and surprise.

Variable	Description	N	Mean	Std. dev.	Min	Max
SUS^{Markov}	Suspense based on Markov	8,563	0.0149	0.0410	1.93e-07	0.2121
SUR^{Markov}	Surprise based on Markov	8,563	0.0122	0.0179	0	0.2717
SUS^{odds}	Suspense based on betting odds	8,563	0.0233	0.0256	4.99e-07	0.2601
SUR^{odds}	Surprise based on betting odds	8,563	0.0129	0.0197	0	0.3097

Notes: The table reports summary statistics for the 80 men's singles matches played at Wimbledon between 2009 and 2014, for a total of 8,563 minutes of live tennis.

and standard deviations of suspense and surprise are comparable between models, although both are slightly higher when computed with the odds-derived belief. The correlation between suspense and surprise is 0.36 (0.38) for the Markov chain (betting odds) method. Finally, we observe that suspense and surprise increase over the stages of the tournament.[17]

4.2 Empirical methodology

The type of data used in this study presents two advantages with respect to the estimation methods. First, our panel data allow us to control for time-invariant factors that might jointly affect the audience level by using within-match variation. Those factors might be the stage of the competition, the day of the week, or the quality of the players.[18] Second, as opposed to stadium attendance, short-term TV audience variation is not affected by factors such as supply capacity, gate price, location, or weather conditions (Borland & MacDonald, 2003; Alavy et al., 2010).

Simultaneously, to account for other characteristics that are subject to change *during* a match that might affect the minute-by-minute TV ratings, we use several control variables. To begin with, we introduce time dummies that correspond to the elapsed time (in minutes) from the start of the match. This allows us to control for any time-related audience differences in a flexible manner. Because all channels offer their

[17]Descriptive statistics of the main variables at the tournament-stage level are provided in the Appendix A.3, Part II.

[18]Rodríguez et al. (2015) illustrate the importance of including a large set of control variables when examining aggregate TV audience measures. As their dependent variable is the average TV audience over the length of the program (they analyze professional cycling races), they also control for, among other things, calendar variables, the scheduling of rival channels, and the competitive balance before the race.

best TV content in the evening hours, potentially turning viewers away from tennis, we introduce a *primetime* dummy, which takes the value of one for all the minutes after 8:00 p.m. and zero otherwise.[19] Intuitively, the coefficient for *primetime* should be negative.

Due to the popular news program on SRF1 (the first national channel), the TV audience level of SRF2 might drop as viewers switch to the newscast. Therefore, we construct the *news* indicator variable, which equals one for any minute between 5:58–6:06 p.m. and between 7:28–7:56 p.m., indicating the first short newscast (6:00–6:05 p.m.) and the following long one (7:30–7:55 p.m.), respectively, and zero otherwise. Intuitively, the coefficient for *news* should be negative. Other programs are not likely to systematically affect the audience variation, particularly because the matches take place at various times of the day, from Monday through Sunday. As Alavy et al. (2010) note, minor variations in audience might be due to channel hoppers. Nonetheless, these authors argue that any moment providing high entertainment should attract even channel hoppers and keep them watching, thus reducing the noise from their behavior.

To allow the players to rest and switch sides of the court, small breaks take place after odd-numbered games and between sets. As these breaks may cause viewers to briefly stop watching, we introduce the *pause* indicator variable, which equals one during the break and zero otherwise.[20] The *pause* variable is needed only for the regression based on betting odds. By construction, the suspense and surprise measures computed with the Markov model are missing during the break because the score does not change.

Regarding the estimation model, as the Wooldridge test evidenced autocorrelation

[19]The results are robust to several alternative definitions of *primetime*, e.g., 6:00 p.m. and 7:00 p.m.

[20]TV broadcasters also use certain breaks for showing commercials. However, we do not insert an *advertising* dummy because the *pause* dummy also captures the effect of viewers switching channels to skip the advertising. To identify time-outs for advertising, we examine the channel content description for SRF2 and SRFinfo, which lists all the programs broadcast and their exact start and end times. Commercial breaks occur on average on a 35-minute basis for all non-finals matches and on a 20-minute basis for the finals. As SRF2 is a national public channel, the amount of advertising is limited by law.

in the audience ratings, we add four lags of *audience* into the regression equation (Wooldridge, 2010). We therefore estimate a dynamic model by applying the Arellano-Bond generalized method of moments (GMM) technique, as the GMM estimators are consistent estimators for dynamic panels (Arellano & Bover, 1995).
Our baseline regression model is specified as follows:

$$\begin{aligned} audience_{i,t} = \alpha_0 &+ \beta_1 \cdot SUS_{i,t} + \beta_2 \cdot SUR_{i,t} \\ &+ \beta_3 \cdot primetime_{i,t} + \beta_4 \cdot news_{i,t} + \beta_5 \cdot pause_{i,t} \\ &+ \sum_{k=1}^{4} \theta_k \cdot audience_{i,t-k} + time\ dummies + \upsilon_i + u_{i,t}\ , \end{aligned} \tag{9}$$

where the subscripts i and t denote match and minute, respectively, and υ_i denote the match fixed effects. The dependent variable $audience_{i,t}$ represents the match i's average live TV audience level at minute t, $\sum_{k=1}^{4} audience_{i,t-k}$ the four lags of the dependent variable, and *time dummies* represent the minutes elapsed since the beginning of the match. The coefficients of interest are β_1 for suspense (SUS) and β_2 for surprise (SUR). For all regressions, we use robust standard errors that are clustered at the match level. Diagnostic and robustness checks are discussed in Subsections 5.2 and 5.3.

5 Empirical results

5.1 Univariate evidence

Before we turn to the estimation results of Eq. (9), we report the results of a univariate analysis of our data. Due to programming conflicts, the broadcasting of tennis is sometimes switched from SRFinfo to SRF2 or vice versa. These switches occur independently of the standing of the current match and thus independently of the moment's level of suspense and surprise. Fig. 2 depicts such an example from a 2012

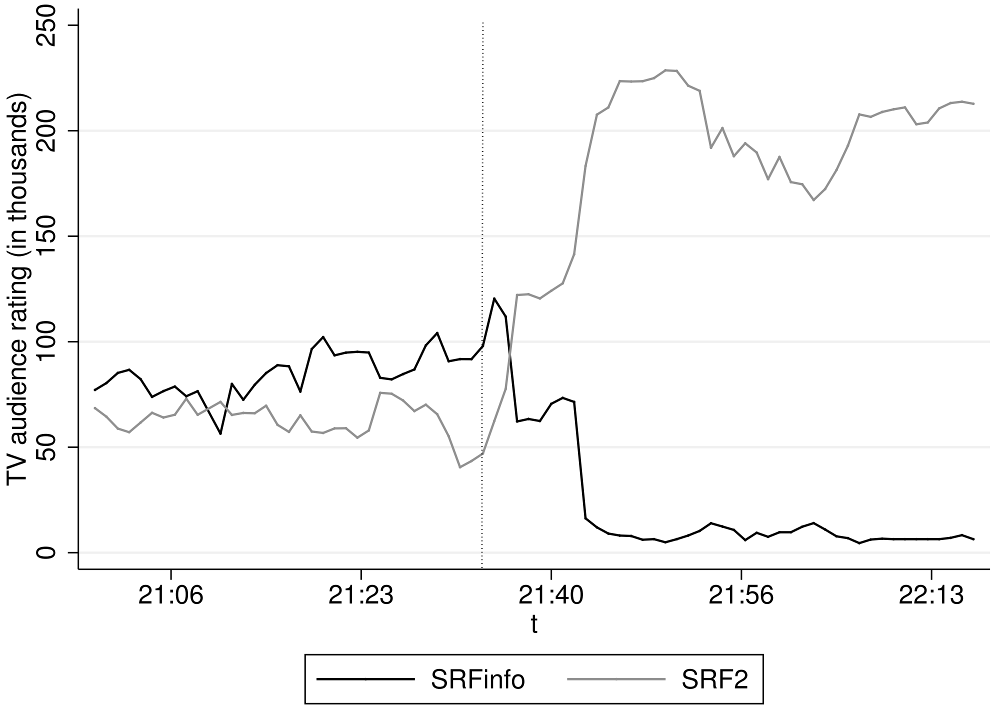


Figure 2
Displayed is the evolution of the TV audience ratings around an exogenous channel switch. The match between Federer and Benneteau was initially broadcasted on SRFinfo and then switched to SRF2 at 9:35 p.m, as indicated by the vertical dotted line. For about three minutes, both channels showed the same content. The SRF2 and the SRFinfo are two free Swiss national channels.

third-round match between Roger Federer and Julien Benneteau, initially broadcast on SRFinfo and then switched to SRF2 at 9:35 p.m. Clearly, the audience switched channels to follow the match, indicating that the audience was actively following.

However, the relevant question is whether the size of the audience change depends on the levels of suspense and surprise during the previous minutes. To answer this question, we identify eight switches that occurred during the broadcast of matches in our sample. For each switch, we can define the precise moment when the live broadcast was interrupted and then continued on the other channel. We measure the jump in ratings on the channel where the broadcast is continued by taking the difference between the average 15-minute ratings preceding and following the switch. We sum SUS^{Markov} and SUR^{Markov} over 15 minutes before the switch, sum the two totals, and compute the median. We create two groups based on whether suspense and surprise were above or below the median. Intuitively, if the 15 minutes before the switch offered above-median (below-median) entertainment, this should be reflected in a larger (smaller) jump in the TV audience on the post-switch channel.

We use the Wilcoxon rank sum test to compare differences in TV audience variation after a broadcaster-initiated channel switch between groups with low and high suspense and surprise. We find that the group with below-median suspense and surprise shows an increase in TV audience of 29,900 viewers, while the group with above-median suspense and surprise shows an increase in TV audience of 99,450 viewers ($z = 2.309$, $p < 0.05$).[21] Overall, this finding provides not only evidence for an active audience assumption but also preliminary and suggestive evidence supporting the hypothesis that a TV program's level of entertainment, measured in terms of suspense and surprise, affects its performance in terms of audience.

5.2 Regression analysis

Table 3 reports regression estimates for the effects of suspense and surprise on TV audience ratings. The Arellano-Bond tests for autoregressive errors yield the expected results (Arellano & Bover, 1995): autocorrelation exists in the first lag but not in the second. Additionally, the Sargan tests of over-identifying restrictions support the null hypothesis that the instruments are valid.

Overall, the results are in line with the hypothesis that moments that offer more suspense and surprise generate more entertainment demand. Columns (1)–(3) show the results for suspense and surprise measures based on the Markov chain model. Columns (1) and (2) demonstrate that suspense and surprise each have a positive effect on TV ratings. Column (3) shows that even when we include both variables together, suspense and surprise remain significant predictors for the TV audience. This result indicates that suspense and surprise are both driving forces behind media

[21] The results are similar when using suspense and surprise from the betting odds method.

Table 3
The effect of suspense and surprise on TV audience.

	Dependent variable: *audience*					
	Markov			Betting odds		
	(1)	(2)	(3)	(4)	(5)	(6)
SUS	46.542***		30.803***	56.867***		44.512***
	(12.883)		(10.122)	(16.824)		(15.472)
SUR		172.861***	146.984***		144.051***	95.404***
		(28.570)	(29.401)		(30.355)	(31.081)
primetime	0.463	0.870	1.064	0.871	0.904	1.167
	(0.412)	(0.687)	(0.660)	(0.631)	(0.751)	(0.770)
news	-1.692***	-1.511***	-1.448**	-1.719***	-1.775***	-1.603**
	(0.531)	(0.590)	(0.614)	(0.614)	(0.659)	(0.665)
pause				-4.138***	-3.761***	-3.918***
				(0.842)	(0.804)	(0.799)
Audience lags	Yes	Yes	Yes	Yes	Yes	Yes
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Match fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
N	8,332	8,332	8,332	8,387	8,387	8,387
AR(1) z-test	-4.59	-4.62	-4.62	-4.44	-4.54	-4.49
$Pr > z$	0.000	0.000	0.000	0.000	0.000	0.000
AR(2) z-test	0.35	0.11	-0.05	-0.17	-0.34	0.03
$Pr > z$	0.724	0.910	0.958	0.863	0.734	0.972
Sargan test $Pr > \chi^2$	1.000	1.000	1.000	1.000	1.000	1.000

Notes: The table reports the results of panel regressions using the Arellano-Bond GMM estimation method. The dependent variable is the TV audience rating (in thousands). The main independent variables, suspense and surprise, are derived from the Markov model (Markov) and in-play betting odds (Betting odds). All estimations also include a constant (not reported). Time dummies correspond to the elapsed time (in minutes) from the start of the match. The data are at the minute-level and includes 80 men's singles matches played at Wimbledon that were broadcast live on SRF2 and SRFinfo from 2009 to 2014. Robust standard errors that have been adjusted for clustering at the match level are provided in parentheses. In all models, *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

entertainment demand. On average, a one standard deviation increase in suspense raises the audience by approximately 1,260 viewers per minute, whereas a one standard deviation increase in surprise raises the audience by approximately 2,630 viewers per minute. As an illustration, the combined effect of a one standard deviation increase in both suspense *and* surprise results in an increase of approximately 3,900 viewers per minute, corresponding to a 3.65% minute-level increase (with respect to an average audience per match of 106,000).

Columns (4)–(6) present the results for suspense and surprise measures based on the in-play betting odds. In all specifications, suspense and surprise have a positive and significant coefficient. According to column (6), on average, a one standard deviation increase in suspense raises the audience by approximately 1,140 viewers per minute, whereas a one standard deviation increase in surprise raises the audience by approximately 1,880 viewers per minute. As an illustration, the combined effect of a one standard deviation increase in suspense *and* surprise results in an increase of approximately 3,000 viewers per minute, corresponding to a 2.83% minute-level increase (with respect to an average match audience of 106,000). Thus, the economic effects estimated with both methods are comparable.

Moreover, surprise appears to be more important than suspense in entertainment demand. The t-statistic of the equality of the estimated coefficients for suspense and surprise is strongly significant for the Markov method ($\chi^2 = 12.39$, $Pr > \chi^2 = 0.000$) and marginally insignificant for the betting odds ($\chi^2 = 1.76$, $Pr > \chi^2 = 0.184$). Depending on the model used to estimate the audience's beliefs, the estimated coefficients for surprise are from two to five times larger than those for suspense.

The coefficients of the control variables mostly have the signs that are expected. The coefficient for the first lag of audience is always very close to one, suggesting some short-term inertia in viewership, whereas the other three lags are small in magnitude and mostly significant. The coefficient for *primetime* is always positive but not statistically significant; therefore, there is no clear evidence of an increase in the competitive mix of TV programs offered during the evening. The coefficient for *news* is always negative and significant at the 5% level, thus confirming our hypothesis that the daily news on SRF1 may attract some viewers away from a tennis match. The coefficient for *pause* is always significantly negative, indicating that some viewers rapidly switch channels during the pauses (after odd-numbered games or between the sets), possibly to skip the commercials.

To investigate the question regarding whether suspense and surprise become more important as a match progresses, we introduce two interaction variables. First, we investigate whether the effects of suspense and surprise are different among sets (*set*): playing more sets leads to slower information revelation, which might generate additional entertainment value from suspense and surprise (Ely et al., 2015). Second, we investigate whether the effect of suspense and surprise is different between the third, fourth, and fifth sets jointly ($late\_set = 1$) and the group constituted by the first two sets ($late\_set = 0$): as the earliest that matches can be won is in the third set, the first two sets might be less entertaining.[22]

Table 4 presents the results. Because our interest is in how these two interaction variables influence the effects of suspense and surprise on TV audience ratings, the table reports only the main effects and the coefficients on the interactions, i.e., any incremental impact that these factors have on the audience ratings. Also included in the specifications but not shown in the table are the control variables for the main specification.

Overall, the evidence supports our assumptions: the interaction coefficients for suspense and surprise with *set* (Panel A) are positive and significant in both models. The interaction coefficients are particularly high for surprise. For example, according to column (2), for each unit change in *set* (i.e., each additional set) the slope of the suspense on audience increases by approximately 6,200 viewers and the slope of surprise on audience increases by approximately 16,900 viewers.

The evidence from the second interaction model (Panel B) corroborates the notion that both suspense and surprise generate even more entertainment value during a match's later moments, i.e., when the stakes are higher. Again, the interaction coefficients are particularly high for surprise. For example, according to column (4), for each unit change in $late\_set$ (i.e., being in a potentially decisive set) the slope of suspense on audience increases by approximately 22,400 viewers and the slope of surprise on audience increases by approximately 66,420 viewers. Overall, the coefficient on the interactions appears to be economically non-trivial.

[22] Differently from the fourth and fifth set, the third set can end a match only when a player leads 2–0 in the third set. Thus, in a further test, we use a slightly different definition: the variable $late\_set$ equals one if the set number is the $3rd$ *and* one player had a 2–0 lead, $4th$, or $5th$ set, and zero otherwise. The results are similar.

Table 4
Interaction models.

Panel A: Interaction with *set*

	Dependent variable: *audience*	
	Markov	Betting odds
	(1)	(2)
SUS	5.294*	6.056*
	(2.941)	(3.562)
SUR	31.982*	11.087
	(16.832)	(9.254)
set	-2.911	-2.411
	(2.211)	(1.998)
SUS × *set*	4.241***	6.285**
	(1.443)	(3.142)
SUR × *set*	28.551***	16.950*
	(7.837)	(9.416)
Control variables	Yes	Yes
N	8,332	8,387

Panel B: Interaction with *late_set* (dummy)

	Dependent variable: *audience*	
	Markov	Betting odds
	(3)	(4)
SUS	29.151**	31.155*
	(11.711)	(16.914)
SUR	50.221*	32.355*
	(27.898)	(18.563)
late_set	4.476	5.653
	(3.269)	(4.352)
SUS × *late_set*	13.525**	22.412***
	(5.132)	(7.308)
SUR × *late_set*	82.920***	66.425***
	(21.734)	(23.415)
Control variables	Yes	Yes
N	8,332	8,387

Notes: The table reports the results of panel regressions with interaction terms. The dependent variable is the TV audience ratings (in thousands). The main independent variables, suspense and surprise, are derived from the Markov model (Markov) and in-play betting odds (Betting odds) and interacted with different variables. The variable *set* takes discrete values between 1 and 5. The variable *late_set* equals one if the set number is the $3rd$, $4th$, or $5th$ set, and zero otherwise. The control variables (four audience lags, primetime, news, pause, match fixed effects, and time dummies) and a constant are included but not reported. Robust standard errors that have been adjusted for clustering at the match level are provided in parentheses. In all models, *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

5.3 Robustness checks

Alternative specifications and postestimation tests

In this subsection, we provide the results of further robustness checks. We begin by discussing alternative specifications and our postestimation analyses (untabulated results). First, we estimate a panel regression with one lag of audience on the right hand side of the equation. Although our model contains a lagged dependent variable, Nickell bias should not be an issue, as we work in a “large T, large N” context (Nickell, 1981).[23] Second, we estimate the model without the lagged audience but allowing the error term to be first-order autoregressive ($u_{i,t} = \rho \cdot u_{i,t-1} + \eta_{i,t}$, where $|\rho| < 1$ and $\eta_{i,t}$ is independent and identically distributed with a mean of 0 and a variance of σ_η^2). Finally, as the dependent variable (*audience*) is a nonnegative integer, we also use the panel Poisson regression and the panel Negative Binomial regression, both with one lag of audience on the right hand side (Winkelmann, 2013). Overall, the coefficients of suspense and surprise are always positive and highly statistically significant, with surprise being consistently larger than suspense throughout all the specifications.

Second, we control in different ways for other time effects. We estimate models including a linear and quadratic continuous time trend in the form of the elapsed time during a match and time of the day (at both the minute and hour levels). All the results are robust to these alternative specifications.

Third, we investigate whether our main results are robust to outliers in the dependent variable. We use two procedures for reducing the effects of the tails: in the first, we delete the 5% (or 1%) largest and smallest ratings; in the second, we winsorize the top and bottom 200 (or 100) observations. Regressions that use trimmed or winsorized audience figures produce similar results.

Fourth, we also check the data for multicollinearity. As discussed in Ely et al. (2015), realized suspense and surprise tend to be positively correlated, which conforms to the "intuition that more suspenseful events also generate more surprise" (Ely et al., 2015, p. 245). Although multicollinearity would not reduce the reliability of the model as a whole, it might result in unstable coefficient estimates and wildly inflated standard errors. The overall sample correlation between suspense and surprise is 0.36 (0.38) for the Markov chain (betting odds) method, which is relatively low. Nonetheless, we

[23]"N" represents the number of cross-sectional units and "T" the number of time points.

compute the variance inflation factors (VIF) for all the dependent variables: based on the results of this analysis, we can conclude that multicollinearity is not a problem.[24]

Finally, we test the sensitivity of our results with regard to S_i, i.e., the player's probability of winning a service point. We replicate all analyses using either the historical S at Wimbledon (0.66) or the average S in our sample (0.69) for all players. The results are robust to these alternatives.

Swiss players subgroup

In an experimental study on the effects of suspense on enjoyment, Peterson and Raney (2008) note that the utility from watching a sportscast also depends on the viewer's disposition toward the participants. Disposition theory describes this effect in detail and proposes that "enjoyment derived from witnessing the success and victory of a competing party increases with positive sentiments and decreases with negative sentiments towards that party" (Zillmann, Bryant, & Sapolsky, 1989, p. 162). Because our data come from Swiss households and because 35 of the matches in the sample include Roger Federer or Stan Wawrinka, both of whom are Swiss players, the audience might show an affective disposition toward these players.[25] Notably, the mean audience for these 35 matches is three times higher than for the other 45 matches: 168,000 viewers ($\sigma = 182,140$) versus 56,733 ($\sigma = 39,823$). Although the overall audience level is not important to our analysis, it might nonetheless indicate a different minute-level TV behavior of the audience.

To address disposition theory, we therefore perform another analysis of the TV audience variation for Swiss and non-Swiss players separately. In Table 5, Panel A shows that suspense and surprise are important entertainment factors in both sub-

[24] We check whether the VIF values are below 10, a generally accepted level indicating that multicollinearity exerts no significant impact. The VIF values for all variables in the Markov (betting odds) models range from 1.03 (1.02) to 1.55 (1.64), with a mean value of 1.26 (1.28). For running this test, we exclude the time dummies from the main regression equation specified by Eq. (9), as they all have a VIF of 1.00 and may deflate the mean VIF (in both models, the mean VIF would be 1.02).

[25] Overall, Federer played in 30 matches and Wawrinka in six; only one match saw them play against one another. Regarding the last three stages of the tournament included in our sample, Federer played 3/6 finals, 3/12 semifinals, and 4/14 quarterfinals, whereas Wawrinka played only one quarterfinal.

Table 5
Robustness checks.

Panel A: Swiss players subgroup				
	Dependent variable: *audience*			
	Markov		Betting odds	
Player group:	Swiss=1	Swiss=0	Swiss=1	Swiss=0
	(1)	(2)	(3)	(4)
SUS	21.736***	11.251***	30.997**	28.496***
	(7.967)	(2.932)	(13.410)	(10.160)
SUR	108.217***	71.021**	70.748**	47.434***
	(36.567)	(29.635)	(31.954)	(15.479)
Control variables	Yes	Yes	Yes	Yes
N	3,664	4,668	3,714	4,673
No. of matches	35	45	35	45
Panel B: Audience by gender				
	Dependent variables: ♀ *audience*, ♂ *audience*			
	Markov		Betting odds	
Audience group:	♀ *audience*	♂ *audience*	♀ *audience*	♂ *audience*
	(5)	(6)	(7)	(8)
SUS	11.633***	21.901***	18.321***	24.765***
	(4.068)	(7.674)	(6.956)	(9.257)
SUR	38.396***	114.715***	24.067**	83.013***
	(12.224)	(22.706)	(12.007)	(21.583)
Control variables	Yes	Yes	Yes	Yes
N	8,332	8,332	8,387	8,387
No. of matches	80	80	80	80

Notes: The table reports the results of panel regressions. In Panel A we distinguish between matches with at least one Swiss player taking part in them (Swiss=1) and no Swiss player (Swiss=0). The dependent variable in Panel A is the TV audience ratings (in thousands). In Panel B, we distinguish the female audience (♀ *audience*) from the male audience (♂ *audience*) for the dependent variable. The main independent variables, suspense and surprise, are derived from the Markov model (Markov) and in-play betting odds (Betting odds). The control variables (four audience lags, primetime, news, pause, match fixed effects, and time dummies) and a constant are included but not reported. Robust standard errors that have been adjusted for clustering at the match level are given in parentheses. In all models, *, **, and *** denote significance at the 10%, 5% and 1% levels respectively.

samples and that surprise has a larger effect, thus confirming our main results. However, it appears that surprise has an even larger effect when Swiss players are on the court. As Federer has regularly been very successful in the Wimbledon Championships tournament—he holds the record for most singles championships won (eight)—the entertainment derived from surprising moments seems to be positively amplified. Finally, the results of a regression with only a subgroup of Federer's matches are almost identical to our main results.

Audience by gender

In a study on the relationship between gender and audience experiences with televised sports, Gantz and Wenner (1991) note that men are more likely than women to become emotionally involved in sports contests and are more responsive while watching. Hence, our results might be driven solely by male viewers. Therefore, we test the validity of the results for male and female viewers separately. In our sample, the male audience (59,308 viewers) is on average higher than the female audience (47,177 viewers). In Table 5, Panel B shows the results for the subsample of female viewers (odd-numbered columns) and the results for the subsample of male viewers (even-numbered columns). The coefficients of suspense and surprise are highly significant and positive for both genders. In particular, the male audience appears to be more responsive to suspense and to surprise, confirming the idea that men are more likely to enjoy the drama and tension involved. Interestingly, columns (5) and (7) show that female viewers are also more markedly responsive to surprising moments than to suspenseful moments. Overall, this evidence suggests that our results are equally valid for male and female viewers.

6 Concluding remarks

6.1 Discussion

When designing entertainment content, decision makers should take into account that both suspense and surprise matter but also that surprise seems to be more important than suspense. In tennis, as suspense and surprise are exogenously determined by the rules and the players, we recognize that it would be difficult to increase either artificially. However, new rules were tested: thus, in 2015, matches without advantage scoring and with sets of first-of-four games were played (CNN, 2015). As we find that suspense and surprise are more important during a match's later moments, such rules targeted at reducing the length of matches might not be a good idea.

Furthermore, our results can be used to evaluate the format of sports competitions. Major League Soccer in the U.S., for example, is a closed league that strongly focuses on high suspense by inducing a fixed number of teams to compete with comparable levels of talent that are enforced through uniform salary caps and extensive revenue redistribution. In contrast, various European soccer leagues allow teams to be more heterogeneous in expenditures on talent and consequently in playing strength. However, relegation and promotion of European clubs based on performance merit ensures that disparities in playing strength cannot exceed certain levels within one league. Thus, the European setting increases the potential for surprise, as underdogs regularly encounter clear favorites and occasionally beat them, but does not ignore the value of suspense. Moreover, parallel European club competitions, such as the Union of European Football Associations (UEFA) Champions League, add additional suspense by matching comparably strong top European clubs.

Most importantly, our findings have implications for the design of entertainment content, particularly where suspense and surprise can be endogenously determined through a rigorous design of their contents, such as movies, TV shows and series, or talent contests. Our methodology can be applied to effectively measure how entertaining an audience perceives the product to be. The challenge remains in measuring people's beliefs. Nonetheless, even for settings in which prediction markets are unavailable or where theoretical modeling is extremely problematic, measurement of beliefs can be feasible.

For example, social media analytic tools can be used to analyze users' posts and comments on social media platforms such as Facebook. Just as tennis fans talk about and want to hear about Wimbledon on Twitter, TV series fans do the same. Alternatively, beliefs might be derived from historical data. For example, to measure the average audience's belief that a particular movie will have a happy ending, one might use as prior belief the historical fraction of movies of a certain genre that have a happy ending and compute suspense and surprise using the actual type of movie ending, i.e., the posterior belief.

Last but not least, the timing and selling of TV advertisements can also be improved. As suspense and surprise enhance entertainment and simultaneously attract human attention, commercials will be more effective following very suspenseful or surprising moments. A larger and more attentive TV audience can be reached, thereby increasing the advertisement's effectiveness. Moreover, the understanding of the effects of suspense and surprise could be translated into higher advertising revenues. For instance, broadcasters might auction advertising slots to companies willing to advertise their products. Hence, it is in the interest of content providers that bidding firms understand the impact of suspense and surprise on the audience: a slot after a very surprising moment could be sold for large amounts. Of course, as more and more entertainment content is available online, including the Wimbledon tournament at *wimbledon.com*, these implications are not restricted to TV commercials.

6.2 Conclusion

Understanding how and when enjoyment from suspense and surprise affect entertainment demand is essential for designing entertainment content. Our paper provides evidence that both suspense and surprise drive entertainment demand, for both men and women, and become increasingly important over time. Our results also suggest that surprise matters more than suspense. We draw this inference by estimating audience beliefs and relating them to high-frequency TV audience figures in the real-world setting of the Wimbledon Championships tennis tournament.

Based on the results of this paper, we discussed important implications for sports and for the design of entertainment content in general. However, we recognize that the issue of suspense and surprise is multifaceted. Specific time patterns and combinations of suspense and surprise may produce different levels of entertainment. For example, an audience might better “tolerate” a boring moment when it follows (or precedes) a very entertaining moment; conversely, an audience might feel anxious if there are long periods with too much suspense or surprise. Because of data limitations, we were not able to precisely distinguish the effects of suspense and surprise from other factors that may also determine entertainment demand. Mood management, escapism, or learning motivations, for example, might be relevant in contexts other than sports. Ideally, detailed data at the individual level might also be used to thoroughly examine individual reactions to various levels of suspense and surprise. We leave these important subjects for future research.

A Appendix

A.1 Tennis

This appendix introduces the basic rules of tennis (Part I) and explains how we compute the probability of winning a point on service (Part II).

Part I Player 1 begins the match by serving in the first game of the first set. Player 1 wins a *point* (sometimes referred as "point game") if player 2 cannot return the ball. A *game* is won when one player wins four points with a two-point difference, or when there is a two-point difference after a deuce, i.e., a score of 40–40 (3 points to 3 points in a game). The players alternate serving every game, and they change ends after every odd-numbered game. A *set* is won when a player either wins six games with a two game difference, or, in the case of a tie-break when the score for one player is 7:6. The *tie-break* begins when the game score is tied at 6:6, and is played until one player wins seven points with a two-point difference, or until there is a two-point difference when the point score is 6–6. At Wimbledon, a tennis match is played as the best of five sets (instead of three), meaning that a player winning three sets wins the match. The fifth set does not have a tie-break; the set is won when one player has won six games and two games more than the other. For further details on the rules of tennis, please consult the official website of the International Tennis Federation (ITF): http://www.itftennis.com/media/220771/220771.pdf.

Part II To compute the probability of winning a point on service for each player in our sample using historical data, we download the necessary statistics from the ATP World Tour website, where official player-level data are available. For all the 160 server-match combinations in our sample (80 matches, two players per match), we download the player's average "*% 1st service in*", "*% of points won if 1st service in*", and "*% of points won if 2nd service in*" statistics from the previous year on grass courts. For players with no history on grass courts for the previous year, we use statistics from two years earlier or, when also unavailable, from the previous year but on hard courts. Then, we apply the formula provided by Klaassen and Magnus (2014, p. 75):

$$\begin{aligned} S_i = & \ (\%\ 1^{st}\ \textit{services in}) \cdot (\%\ \textit{of points won if}\ 1^{st}\ \textit{service in}) \\ & + (\%\ 1^{st}\ \textit{services not in}) \cdot (\%\ \textit{of points won if}\ 2^{nd}\ \textit{service in}) \end{aligned} \tag{A.1}$$

The first part of the formula reflects the probability of winning the point on the first serve, whereas the second part reflects the probability of winning the point on the second serve when the first serve is faulted. Thus, S_i reflects the fact that a player can win a point on either the first *or* second service. As an illustration, we report a numerical example provided in Magnus and Klaassen (1999) using Wimbledon data on men's singles from 1992 to 1995: $S_i = 0.587 * 0.777 + (1 - 0.587) * 0.518 = 0.456 + 0.214 = 0.67$. In our sample, $S_i = 0.69$ on average: Federer has the highest S_i (0.787 in 2009), whereas Albert Ramos-Vinolas has the lowest (0.483 in 2012). Overall, the advantage of this procedure over using a fixed S for all matches is that we use the same statistics that are also readily available to gamblers.

A.2 Mediapulse TV panel

This appendix provides details on the Mediapulse TV panel. Every household in the panel is given a small measuring device that is connected to all TV sets in the house or apartment. This field-tested device is used in almost 25 countries worldwide. It collects audience information every second. Mediapulse then aggregates these data for the entire panel and saves it at the minute-by-minute level. For example, the TV audience corresponding to the minute 15:23:00 indicates the average audience between 15:23:00 and 15:23:59.

Extrapolation: the reporting samples of a panel are extrapolated to the population (universe) estimates, which allows the results to be representative of the respective population, i.e., of the TV panel target audience (Kuonen & Hulliger, 2013). For extrapolating, Mediapulse uses actual population figures from the Swiss Federal Statistical Office. The quota attributes used to determine the appearance of a household in the panel are language area, canton (a member state of Switzerland), district (a region of a canton), household size, presence of children in the household, and age of the head of household. To certify that the panel conforms to international quality standards, it is subject to an annual external verification.

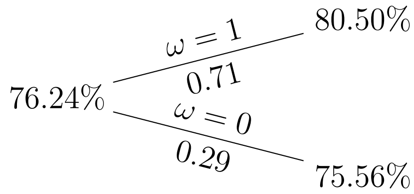
Changes in the panel: over the 2009–2014 period, two changes in the measuring system have occurred. First, beginning in 2010, daily weighting was introduced, and the use of Teletext was considered to be TV watching. Second, before 2013, a panel of 1,918 households was randomly recruited by the phone. This universe comprised all households with at least an installed TV. However, after 1st January, 2013, the panel was recruited by either the phone or mail. It now contains 2,000 households, at least 1,870 of which must provide data daily. For both panels, households watching TV exclusively from a computer are not included. Mediapulse informed us that for an audience analysis *within* a program, like a tennis match, none of these changes has had an impact on the size of the variation in the measured audiences.

A.3 Markov model and further descriptive statistics

This appendix illustrates how to compute SUS^{Markov} and SUR^{Markov} (Part I) and provides further detailed descriptive statistics for the main variables in our model (Part II).

Part I Using a fictitious example, we illustrate how to compute SUS^{Markov} and SUR^{Markov}, i.e., suspense and surprise based on a finite binary Markov chain. An almost identical procedure applies to computing SUS^{odds} and SUR^{odds}. Player W is playing against player L. After the $63rd$ point, the match is tied at one set apiece, with player W ahead five games to four (5–4) in the set, and 40 points to zero (40–0) in the game. Therefore, player W is serving to win the game and, in so doing, also the set (the expression used is: "Player W has three set points"). From his historical serving records on grass, we compute server W's probability of winning the *next point*, $S_W = 0.71$ (see Part II of Appendix A.1). As the situation of the players in a tennis match is always symmetrical, player L's probability of winning on a return point is 0.29.

Using this information, the model predicts that player W's match winning probability is 76.24% (i.e., the current belief μ_p). If he wins the point ($\omega = 1$), the posterior belief would rise to 80.50% (+4.26%). However, as there is a 29% chance that he will lose next point ($\omega = 0$), that loss would bring the posterior belief down to 75.56% (−0.68%). The probability transition from point 63 to point 64 is:



where 80.50% and 75.56% are the posterior beliefs. The possible size of the update in the beliefs is correctly asymmetric (+4.26% vs. −0.68%), as player W would still have two set points left to play even if he loses this serve point. Finally, if player W indeed loses the $64th$ point ($\omega = 0$), the posterior belief would drop to 75.56%, making 80.50% the counterfactual probability. Using Eq. (2), we compute suspense for the 63rd point ($p = 63$) as the standard deviation of the next point's beliefs:

$$SUS_{63}^{Markov} = [0.71 \cdot (0.805 - 0.7624)^2 + 0.29 \cdot (0.7556 - 0.7624)^2]^{1/2} = 0.036.$$

Similarly, using Eq. (3) we compute surprise for the 64th point ($p = 64$) as the absolute value of change in beliefs from point 63 to 64:

$$SUR_{64}^{Markov} = |0.7556 - 0.7624| = 0.0068.$$

Thus, the model predicts a suspense of 0.036 and a surprise of 0.0068.

Part II

Table A.1
Descriptive statistics (mean) of the TV audience, suspense, and surprise by tournament stage.

			Markov		Betting odds	
Tournament stage	N	*audience*	*SUS*	*SUR*	*SUS*	*SUR*
$1st$ stage	12	44.0	0.0064	0.0061	0.0131	0.0093
		(29.23)	(0.0121)	(0.0110)	(0.0231)	(0.0175)
$2nd$ stage	14	55.2	0.0096	0.0078	0.0173	0.0098
		(44.35)	(0.0208)	(0.0118)	(0.0362)	(0.0179)
$3rd$ stage	12	75.8	0.0103	0.0084	0.0193	0.0125
		(45.16)	(0.0407)	(0.0175)	(0.0431)	(0.0193)
$4th$ stage	10	83.2	0.01284	0.0096	0.0216	0.0127
		(84.85)	(0.0413)	(0.0177)	(0.0432)	(0.0198)
Quarterfinal	14	87.4	0.0193	0.0154	0.0281	0.0138
		(59.84)	(0.0477)	(0.0195)	(0.0475)	(0.0204)
Semifinal	12	89.1	0.0209	0.0162	0.0320	0.0162
		(54.83)	(0.0502)	(0.0205)	(0.0506)	(0.0205)
Final	6	362.8	0.0224	0.0199	0.0348	0.0167
		(252.38)	(0.0527)	(0.0206)	(0.0582)	(0.0218)

Notes: Displayed are summary statistics for the TV audience ratings (in thousand), suspense, and surprise for the 80 men's singles matches played at Wimbledon between 2009 and 2014, for a total of 8,563 minutes of live tennis. Standard deviations are provided in parentheses.

References

Alavy, K., Gaskell, A., Leach, S., & Szymanski, S. (2010). On the edge of your seat: Demand for football on television and the uncertainty of outcome hypothesis. *International Journal of Sport Finance*, *5*(2), 75–95.

Alwitt, L. F. (2002). Suspense and advertising responses. *Journal of Consumer Psychology*, *12*(1), 35–49.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, *68*(1), 29–51.

Baldi, P., & Itti, L. (2010). Of bits and wows: A Bayesian theory of surprise with applications to attention. *Neural Networks*, *23*(5), 649–666.

Barnett, T., & Clarke, S. R. (2005). Combining player statistics to predict outcomes of tennis matches. *IMA Journal of Management Mathematics*, *16*(2), 113–120.

Borland, J., & MacDonald, R. (2003). Demand for sport. *Oxford Review of Economic Policy*, *19*(4), 478–502.

Bryant, J., Rockwell, S. C., & Owens, J. W. (1994). Buzzer beaters and barn burners: The effects on enjoyment of watching the game go down to the wire. *Journal of Sport & Social Issues*, *18*(4), 326–339.

CNN. (2015). *Fast4: Tennis' answer to cricket's twenty20.* Retrieved 2016-05-13, from http://edition.cnn.com/2015/01/12/tennis/fast4-new-tennis-format/

Comisky, P., & Bryant, J. (1982). Factors involved in generating suspense. *Human Communication Research*, *9*(1), 49–58.

Croxson, K., & Reade, J. (2014). Information and efficiency: Goal arrival in soccer betting. *The Economic Journal*, *124*(575), 62–91.

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, *123*(1), 215–260.

Forrest, D., Simmons, R., & Buraimo, B. (2005). Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy*, *52*(4), 641–661.

Gantz, W., & Wenner, L. A. (1991). Men, women, and sports: Audience experiences and effects. *Journal of Broadcasting & Electronic Media*, *35*(2), 233–243.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *The Journal of Finance*, *46*(1), 179–207.

Horstmann, G. (2015). The surprise–attention link: A review. *Annals of the New York Academy of Sciences*, *1339*(1), 106–115.

Hutchins, B. (2014). *Game, Set, Cash! Inside the secret world of international tennis trading.* Black Inc.

Itti, L., & Baldi, P. (2009). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306.

Klaassen, F., & Magnus, J. R. (2014). *Analyzing Wimbledon: The power of statistics.* Oxford University Press, USA.

Kuonen, D., & Hulliger, B. (2013). *Evaluation of the new Mediapulse television panel with respect to its suitability for local television channels* (Tech. Rep.). Mediapulse AG.

Liu, Y. (2001). Random walks in tennis. *Missouri Journal of Mathematical Sciences*, *13*(3).

Magnus, J. R., & Klaassen, F. J. (1999). On the advantage of serving first in a tennis set: Four years at Wimbledon. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *48*(2), 247–256.

Newton, P. K., & Aslam, K. (2009). Monte Carlo tennis: A stochastic Markov chain model. *Journal of Quantitative Analysis in Sports*, *5*(3).

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica*, *49*(6), 1417-1426.

Olson, J., & Stone, D. F. (2014). Suspense-optimal college football play-offs. *Journal of Sports Economics*, *15*(5), 519–540.

O'Malley, A. J. (2008). Probability formulas and statistical analysis in tennis. *Journal of Quantitative Analysis in Sports*, *4*(2).

Peterson, E. M., & Raney, A. A. (2008). Reconceptualizing and reexamining suspense as a predictor of mediated sports enjoyment. *Journal of Broadcasting & Electronic Media*, *52*(4), 544–562.

Rodríguez, C., Pérez, L., Puente, V., & Rodríguez, P. (2015). The determinants of television audience for professional cycling: The case of Spain. *Journal of Sports Economics*, *16*(1), 26–58.

Su-lin, G., Tuggle, C. A., Mitrook, M. A., Coussement, S. H., & Zillmann, D. (1997). The thrill of a close game: Who enjoys it and who doesn't? *Journal of Sport & Social Issues*, *21*(1), 53–64.

Vorderer, P., Klimmt, C., & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory*, *14*(4), 388–408.

Vorderer, P., Wulff, H. J., & Friedrichsen, M. (2013). *Suspense: Conceptualizations, theoretical analyses, and empirical explorations.* Routledge.

Walker, M., Wooders, J., & Amir, R. (2011). Equilibrium play in matches: Binary Markov games. *Games and Economic Behavior*, *71*(2), 487–502.

Winkelmann, R. (2013). *Econometric analysis of count data.* Springer Science & Business Media.

Wolfers, J., & Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities.* (Working paper)

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data.* MIT press.

Zillmann, D. (1991). The logic of suspense and mystery. In *Responding to the screen: Reception and reaction processes* (Vol. 7, pp. 281–303).

Zillmann, D. (1996). The psychology of suspense in dramatic exposition. *Suspense: Conceptualizations, theoretical analyses, and empirical explorations*, 199–231.

Zillmann, D., Bryant, J., & Sapolsky, B. S. (1989). Enjoyment from sports spectatorship. In *Sports, games, and play: Social and psychological viewpoints* (Vol. 2, pp. 241–278).

A.2

When do professionals play minimax? Evidence from the tennis court

Paolo Bizzozero, Raphael Flepp, Egon Franck

Abstract

This article provides new empirical evidence on von Neumann's minimax theorem. We use tennis as a natural laboratory where professional athletes repeatedly play a two-person zero-sum game: the tennis serve. Using detailed data from 80 men's singles matches between 2009 and 2014 at the Wimbledon Championships tournament, we show that the results are consistent with only one prediction of the minimax theorem: players' choices are serially independent, but their winning probabilities are not statistically identical across strategies. Most importantly, we investigate under what conditions players deviate more from minimax play. In particular, we find that top servers facing non-top receivers typically do not play near the predictions of minimax but rather win more. Our findings reveal that the relative skill level of the players in a match is an important factor to consider when analyzing the players' behavior.

JEL Classification: C72
Keywords: Minimax, Tennis

1 Introduction

In 1928, John von Neumann published the minimax theorem, a formalization of a two-person zero-sum simultaneous mixed-strategy game. The minimax theorem posits that each player minimizes the maximum payoff possible for the other player. As the game is zero-sum, i.e., the payoff of one player is the inverse of that of his or her opponent, each player is also maximizing his or her own minimum payoff. Most important, the minimax theorem provides two testable predictions about how players should behave in strategic situations: (1) the players' winning probabilities are statistically identical across choices, and (2) the players' choices are random. Accordingly, by playing minimax, a player ensures that his or her opponent (1) will not profit from any available pure strategy and (2) will not anticipate his or her actual choice.

O'Neill (1987) argued for the importance of testing the empirical validity of the minimax theorem because many economic models are based on it or its generalization to n-player non-zero-sum games, the Nash equilibrium. Over the last two decades, several laboratory experiments have tested the empirical validity of the minimax theorem. In general, the subjects of experimental studies who play simple games, such as matching pennies or card games, do not behave as predicted by the minimax theorem: their win rates across various strategies are not similar and their choices are not serially independent (e.g., Rapoport & Boebel, 1992; Mookherjee & Sopher, 1997; Binmore, Swierzbinski, & Proulx, 2001; Shachat, 2002; Rosenthal, Shachat, & Walker, 2003; Levitt, List, & Reiley, 2010).

Sports offer a natural setting that solves some of the issues affecting experimental studies, which have primarily used students as study participants. First, as rewards are well defined and clear—e.g., winning in a competition—players are well motivated and incentivized. Second, because professional players dedicate their lives to the practice of a certain sport, they are experienced in playing strategically. Third, professional players always choose actions that maximize their prospects for victory;

thus, researchers do not need to make assumptions about the player's utility function for the monetary payoffs.

Walker and Wooders (2001) and Hsu, Huang, and Tang (2007) use tennis as a natural laboratory where professional athletes play a two-person zero-sum simultaneous mixed-strategy game, the tennis serve. They model the serve game as a simple 2×2 game (both players can play left or right) and assess the empirical validity of the two minimax predictions. Walker and Wooders (2001) analyze 10 men's matches and find that the players' behavior is consistent only with prediction (1), i.e., that the winning probabilities across choices are equal, whereas Hsu et al. (2007) analyze 10 different men's matches and find that the players' behavior is consistent with prediction (1) and prediction (2), i.e., that players' choices are random.

Despite their somewhat different results, a common property of Walker and Wooders (2001) and Hsu et al. (2007) is that both focus on testing whether the observed behavior of professional athletes is "close enough" to the theoretical predictions to be considered consistent with minimax. However, these previous studies do not address the important question of *when* (under what conditions) players deviate more or deviate less from the minimax predictions. Such tests are important because the players' behavior in a real setting may differ under, for example, stress or fatigue, or when playing against opponents with different skills. In our article, we address exactly this open question.

To do so, we begin by re-evaluating the studies of Walker and Wooders (2001) and Hsu et al. (2007) but using a larger and more heterogeneous sample. Our full sample results stand in contrast to previous studies, showing that the players' behavior is consistent only with prediction (2), i.e., that players' choices are random. Then, using a number of subsamples that cover various potential influencing factors, we investigate under what conditions tennis players deviate more or less from both minimax predictions. Most importantly, our results show a tendency toward minimax play in matches between right-handed players, in the later sets of a match, in shorter matches, and especially in unbalanced matches. After splitting the matches according to the skill level of the server and receiver, we find that top servers' behavior is less consistent with minimax play when facing non-top receivers. While this result appears surprising at first glance, we can show that it is rational for top players to deviate: top players would decrease by 18% their likelihood of winning a match by playing closer to minimax against non-top players.

The remainder of this article is organized as follows. Section 2 reviews the applicable literature. Section 3 describes the setting and the hypotheses. Section 4 presents the data and outlines our empirical methodology. Section 5 presents the results. Section 6 discusses the results and concludes the article.

2 Literature review

The empirical literature on minimax can be broadly divided into experimental and field studies, with experimental studies predominating. O'Neill (1987) was the first of a series of original laboratory tests of the minimax predictions. O'Neill matched 25 pairs of students, who then played a card-matching game more than 100 times in succession, each having four strategies to play. O'Neill's results at the aggregate level are very close to minimax play. The results at the aggregate level refer to the results of the minimax tests for all the games *jointly* (e.g., for all pairs of players playing several card games), whereas the results at the individual level refer to the minimax tests for a single game *individually* (e.g., for two players playing a card game).[1]

Later studies revisited O'Neill's card game experiment by proposing design refinements and using different experimental participants; however, these experiments generally failed to provide consistent evidence supporting minimax. Rapoport and Boebel (1992) had 20 students playing a slightly more cognitively demanding card

[1]Section 4 provides further details on this important distinction between the individual and aggregate (or joint) levels.

game, giving higher rewards to increase the participants' motivation, but they still reject minimax play at both the individual and aggregate levels. In Mookherjee and Sopher (1994), two groups of 10 pairs each played a matching penny game, during which the experimenters revealed only to one group of players all of the opponent group's past choices and earnings. The informed subjects played significantly differently from the non-informed subjects, for example by playing somewhat closer to prediction (2). Mookherjee and Sopher (1997), studying the dynamic pattern of play of 20 students playing a table game, also refuted minimax.

To overcome some of the problems that adversely affected previous experimental studies, Shachat (2002) proposes a new method for eliciting mixed strategies. Nonetheless, his results suggest that the minimax predictions do not thoroughly describe the behavior of the 60 students who played an adaptation of the O'Neill game. Rosenthal et al. (2003) propose an even simpler design and, using 40 pairs of students as participants, conclude that although the aggregate choice frequencies are somewhat close to equilibrium, there is excessive serial correlation in the choices. Binmore, Swierzbinski, and Proulx's (2001) experiment gives participants both better incentives and sufficient time for learning, and they find that the players' choices are close to being independent. Binmore et al. underline the importance of using the so-called evolutive interpretation of equilibria, a theory positing that people come closer to equilibrium by undergoing a trial-and-error adjustment process.

To date, field data from professional sports have provided stronger support for the empirical validity of the minimax predictions than have laboratory experiments. Walker and Wooders (2001) and Hsu et al. (2007) focus on the tennis serve: in this game, the server aims at maximizing the expected probability of scoring a point on serve, and the receiver aims at minimizing it. Walker and Wooders' (2001) original contribution, the first that uses field data to test minimax, analyzes the behavior of 13 top players at 10 Grand Slam men's singles matches (mostly finals) between

1974 and 1997.[2] Using point-by-point data collected manually from videotapes, they find supporting evidence for prediction (1), i.e., that the winning probabilities across choices are equal, but not for prediction (2), i.e., that players' choices are random. Hsu et al. (2007) analyze 10 men's, nine women's, and eight juniors' singles matches between 1980 and 2003 and find supporting evidence not only for prediction (1) but also for prediction (2). However, Kovash and Levitt (2009) argue that the tests used in these two tennis studies have low power to reject the null hypothesis because their samples are too small.

Complementing the tennis studies, Palacios-Huerta (2003), Chiappori, Levitt, and Groseclose (2002), and Dohmen and Sonnabend (2016) focus on the penalty kicking between the most experienced kickers and goalkeepers in professional European soccer, generally finding support for both minimax predictions. However, penalty kicks are taken very infrequently, resulting in a bias toward less serial correlation in the players' choices, and often not against the same goalkeepers, resulting in different payoffs. For example, Chiappori et al. (2002) admit that their finding of a lack of serial correlation is "not so surprising since the penalty kicks take place days, weeks, months apart" (Chiappori et al., 2002, p. 1147). A better test for serial correlation is to compare the players' decisions over a relatively short period—e.g., over a match—as we do in tennis.

Palacios-Huerta and Volij (2008) and Levitt et al. (2010) attempt to reconcile the conflicting results of experimental and field studies. To do so, both studies compare the laboratory performance of students to that of professional subjects playing a simple card game. These authors suspect that the strategic skills of the subjects may play an important role in their decision or ability to play according to minimax. They argue that in contrast to the sports athletes in the field studies, students have fewer strategic skills because they are inexperienced at playing mixed-strategy games. In

[2]The Grand Slam tournaments are the four most prestigious tennis competitions worldwide: the Wimbledon Championships (UK), the French Open, the Australian Open, and the US Open.

particular, the researchers select professionals who face strategic situations on a daily basis: Palacios-Huerta and Volij (2008) use soccer players, whereas Levitt et al. (2010) use bridge and poker players.

Although both studies consider similar settings, they surprisingly come to strikingly opposite conclusions: in Palacios-Huerta and Volij (2008), both amateur and professional soccer players, but not students, behave close to minimax, whereas in Levitt et al. (2010), neither students nor professionals do so. Moreover, Wooders (2010) reexamines Palacios-Huerta and Volij's data and comes to the opposite conclusion, that professionals do not play consistently with minimax. Levitt et al. (2010) argue that study subjects, whether students or professionals, might not be able to systematically transfer the strategic skills learned in their professional activities to the laboratory games, as minor differences in the context of the game can have deep behavioral effects.

Overall, the empirical evidence on minimax is mixed. Although the importance of skills is debated by Palacios-Huerta and Volij (2008) and Levitt et al. (2010), no field study has yet investigated the role of the relative skill strengths of sports professionals. Importantly, the skills of professional athletes are generally high but heterogeneous: for example, Roger Federer and Lionel Messi have skills that few other professional players have. Previous field studies have investigated a homogeneous group composed of top-skilled players, i.e., the best tennis players reaching the Grand Slam finals (Walker & Wooders, 2001; Hsu et al., 2007) or the most experienced soccer players competing in the best European soccer leagues (Chiappori et al., 2002; Palacios-Huerta, 2003; Dohmen & Sonnabend, 2016). Therefore, the conclusions of these articles do not generalize to *any* players but only to the most skilled ones.

3 Setting and research hypotheses

Professional sports offer an appropriate setting for studying strategic behavior under natural conditions because the rules of the game are clearly defined and the data are accurately collected by specialized IT firms. In comparison to experiment participants (typically students), professional players are experienced in playing strategically and are highly motivated. Given that a tennis match takes place over a relatively short time, between a fixed pair of players, and with a large number of serve points (typically over 200 in a Grand Slam match), tennis is particularly well suited for studying strategic interactions.

In the Appendix A.1, we briefly describe the rules of tennis that are relevant to this article and display the tennis court. Each half of the court is divided into two equal parts: the left and the right part (when each player faces the court).[3] Each game starts with a player serving from the right court into the cross-court service box. The match continuously alternates between right-court points and left-court points.

Since two players alternate serving from two courts, we distinguish among four different situations for each match: (1) player 1's left-court serves, (2) player 1's right-court serves, (3) player 2's left-court serves, and (4) player 2's right-court serves. Distinguishing among left- and right-court serves is important because serving from different courts requires different abilities and some servers may have preferences for a certain side (Klaassen & Magnus, 2014). We refer to each of the four situations as one *experiment*.[4]

The server decides whether to direct the ball to the receiver's left or right. As in previous tennis studies on minimax, we do not include serves to the receiver's body (center).[5] To increase his chances of reaching and returning the ball, a receiver has to instantly guess, perhaps only subconsciously, the direction of the ball. Consequently,

[3]While the technical tennis terms for the left and right half sides of the court are *ad* and *deuce*, respectively, for ease of reading, this article uses *left* and *right*.

[4]For example, an experiment consists of all of player 1's left-court serves in one match.

[5]Serves directed to the center represent 6.66% of the first serves and 14.55% of all serves in our sample.

he will defend that side more by, for example, being more ready to jump to that side. As an illustration, Brad Gilbert, a former professional player turned tennis coach, advises the receiver to "keep mixing [your] position up so your opponent [the server] has to react to what *you're* doing and not the other way round. [...] One time make your move when they can see you. The next time move just after they make their toss and cannot see you" (Gilbert & Jamison, 2013, p. 107).

Thus, to derive the theoretical minimax predictions, we model each serve as shown in Table 1. Each π denotes the probability of winning a point on serve from the server's perspective, given both players' choices; the first subscript indicates the server's decision to direct the ball to the receiver's left or right, whereas the second subscript indicates the receiver's decision to anticipate and overplay to his own left or right. Initially, as in the previous tennis studies, we assume that this payoff matrix is the same for every point in a given experiment.[6]

Table 1
Payoff matrix.

		Receiver	
		L	R
Server	L	π_{LL}	π_{LR}
	R	π_{RL}	π_{RR}

A unique mixed-strategy equilibrium exists when the following inequalities are satisfied:

$$\pi_{LL} < \pi_{LR}, \quad \pi_{LL} < \pi_{RL}, \quad \pi_{RR} < \pi_{LR}, \text{ and } \quad \pi_{RR} < \pi_{RL}. \tag{1}$$

These inequalities imply that if the receiver correctly anticipates and overplays the side where the server hits the ball, his payoff will be higher and, since each point is a zero-sum game, the server's payoff π will be lower. In tennis, the receiver's decision cannot be observed, and thus, we cannot compute these payoffs. Following Walker and Wooders (2001) and Hsu et al. (2007), we therefore assume that these four conditions

[6]Fig. 1 in Walker and Wooders (2001) provide a hypothetical numerical example of the payoff matrix.

are nonetheless satisfied such that a unique Nash equilibrium in mixed probability strategies exists.

In the first part of this article, we examine the two testable predictions resulting from the existence of a mixed-strategy equilibrium. First, each player should mix his actions to make his opponent's expected payoffs identical across choices. As a consequence, each player's probability of winning a point on serve across choices should be the same (from the same court).

Prediction 1. The server's winning rate for each pure choice $\{L,R\}$ is the same.

The intuition is that each player should mix his choices in such proportions that his opponent cannot exploit him by pursuing any particular pure strategy. Thus, each player's probability of winning a point on serve should be the same whether he plays left or right against his opponent's mixture.[7]

Second, each player should generate a sequence of actions that is unpredictable. As a consequence, each player's serve direction choices should be serially independent.

Prediction 2. The server's choices are serially independent.

The intuition is that each player must be concerned with only the current point, meaning that intertemporal links between points must be absent; at best, a player should randomize over his choices. If a player chooses not to switch his actions often enough or if he switches actions more often than predicted by the theory, his choices will not be serially independent. The same outcome would occur if the player is learning and memorizing his opponent's actions in an attempt to gain a strategic advantage. We include both the first and second services because if we drop the second serves, we cannot observe the full sequence of the servers' choices. Walker and

[7]We emphasize that the first minimax prediction concerns the equality of the serve winning rates, not the equality of the serve direction frequencies. In other words, equalizing the percentage of left and right serves does not equate to equalizing the winning rates. For example, both a player serving 90% of the serve-points to the left and 10% to the right and another serving 50% to both sides may still be playing according to the minimax.

Wooders (2001) and Hsu et al. (2007) focus only on first serves, probably because of the difficulty of hand-collecting the data on second serves.

In the main part of this article, we address the question of *when* tennis professionals play minimax. To simplify the analysis, we initially assumed that the payoffs and thus the equilibrium mixtures vary only across player-court combinations. However, an undetermined set of conditions and factors varying within and across matches can influence a player's likelihood of winning the point (the payoffs π) and thus his strategic behavior. Therefore, we add to the literature by retesting both minimax predictions by distinguishing not only across player-court combinations but also across the different situations detailed below.

First, we distinguish between matches with one left-handed player and matches with only right-handed players—our sample does not include any match between two lefties. Several individual rejections of *Prediction 1* in Walker and Wooders (2001) result from left points played by lefties. When lefties serve left points, they can hit a wide shot onto a righty's backhand, a shot that is very difficult to return. Although righties can use a similar tactic when serving right-court points, they are usually less successful because lefties are much more accustomed to hitting backhand shots. Overall, we expect the players to behave closer to the minimax predictions in righty matches.

We then distinguish between tournament stages. As the Wimbledon tournament is played over seven stages, we split the sample between the first five stages and the last two stages, which include the semifinal and the final. Usually, only the most experienced players qualify for the last tournament stages; thus, we may observe different behaviors there than in the first rounds. Moreover, the semifinal and final subsample permits a comparison of our results with those of previous tennis studies that focus primarily on finals and find supporting evidence for minimax (Walker &

Wooders, 2001; Hsu et al., 2007).[8] Overall, given the results of previous tennis studies, we expect the players to behave closer to the minimax predictions in final and semifinal matches.

Next, we distinguish among sets. Because matches at Wimbledon are played as the best of five sets, we distinguish between the first two sets and the third, fourth, and fifth sets jointly. The earliest that matches can be won is in the third set; thus, in the first two sets, players might feel less pressure and be less nervous. In soccer, Palacios-Huerta (2003) find a lower scoring rate for penalty shots taken in the last 10 minutes of close matches and argue that this might be due to the negative effects of pressure. Moreover, players might learn and improve their strategic skills during the match, which might result in more minimax play (Mookherjee & Sopher, 1994). González-Díaz, Gossner, and Rogers (2012) show that not all players are equally able to increase their performance during important moments of a match. Overall, if the learning effect prevails, we expect the players to behave closer to the minimax predictions in the later sets.

We also divide our sample into two groups according to whether a match has more or fewer total points than the sample median (213 points). Davey, Thorpe, and Williams (2002) show that fatigue plays an important role in tennis, for example leading to a decline in serve accuracy. According to these findings, we expect the players to behave closer to the minimax predictions in shorter matches.

Intuitively, playing against an equally skilled player is not the same as playing against a stronger or weaker player. We retest the minimax predictions across a subsample of balanced and unbalanced matches. For each player, we derive his expected probability of winning the match from the betting odds at the beginning of the match.[9] From the difference between the two players' winning probabilities, we

[8] Due to the low number of finals in our sample (six), we also include the 12 semifinals to have enough observations for the tests.

[9] The inverse of the match-winner betting odds (e.g., 1/1.9=0.52) can be interpreted as the bettors' underlying probabilistic belief in a player winning the match.

know how balanced the match is—a larger difference indicates a more unbalanced match. For example, if at match start the betting market predicts Rafael Nadal to win with a probability of 60% against Andy Murray, the difference is 20%. We define balanced matches as those with the 25% smallest difference and unbalanced matches as those with the 25% largest difference.[10] In general, as there is no literature on this particular subject, we cannot derive a clear prediction.

We further investigate the role of the players' skills in a match. Our previous subsample definition does not distinguish a match between two non-top players from a match between two top players (both are balanced). We now categorize each player with an Association of Tennis Professionals (ATP) ranking above six (the median player ATP ranking) as a "top" player and a player with an ATP ranking below or equal to six as a "non-top" player. We retest both minimax predictions for each of the four possible combinations of top and non-top players. As O'Neill (1987) writes, "Skill would mean that some of the [skilled] subjects could exploit the other's deviation from minimax by estimating the other's individual moves or noticing statistical tendencies" (O'Neill, 1987, p. 2108). Thus, we expect significant differences with respect to minimax play among these four groups.

Finally, in the last part of this article, we ask whether playing minimax leads to winning more points and thus to higher chances of winning the match. For example, after having qualified for the 2015 Wimbledon final against Andy Murray, Roger Federer told a television interviewer that one important factor helping him win was that he "was able to mix it [the serve] well." We also analyze whether the relationship between minimax play and winning depends on the combination of skills of the two players: deviating from minimax play might be rational only if it leads to an increase in a player's likelihood of winning the match.

[10]Our sample encompasses many matches from the early tournament stages, which can be very unbalanced when a top player is participating in them. The average of the odds-implied winning probability difference is 0.67, whereas the median is 0.78. By using 25% as our cutoff, we ensure that the balanced matches are characteristically different from the unbalanced matches.

4 Data and methodology

Detailed point-by-point data at the match level come from IBM, which collects data both directly from the umpires' computers and from analysts who attend the match and manually feed the data into the system. Beyond general information about the match, such as players, court, tournament round, and the start or ending time, the data also contain detailed point-level information on the score, time (in seconds), server, winner, and serve speed. Importantly, the dataset contains information about the serve direction: left, center, or right. In the event of a double fault, when the ball does not land in a valid area on the opponent's side for twice in a row, the serve direction is missing. In addition, we manually collect from the ATP website a player's playing hand and his ATP ranking.[11] Finally, the betting odds for who will win as of the beginning of the match originate from Betfair, one of the largest online betting markets, and are provided by Fracsoft, a data vendor.

Table 2 provides summary statistics.[12] The original IBM sample contains 17,626 point-level observations. After dropping all the serves to the body (2,496 observations) and all the double faults (472 observations), the final sample consists of 14,658 point-level observations from 80 matches. Our sample is heterogeneous, containing an almost equal proportion of matches from any tournament stage over six consecutive years and a total of 58 unique players. A match lasts on average 150 minutes and consists of approximately 220 points, 36 games, and 3.6 sets. Overall, 10 players (17.2% of all players) are left-handed, and the number of matches with one left-handed player is 24 (30% of all matches).

Similar to Walker and Wooders (2001) and Hsu et al. (2007), the average percentage of serve points played to the receiver's left is 55%. On average, the percentage of points won when serving to the left (70.3%) is almost identical to that when serving

[11] The ATP rankings date from the week before the beginning of the tournament. For example, since Wimbledon 2014 started on June 23, the ATP ranking is as of June 17. A lower rating indicates a better player: the No. 1 player is the best player in the world.

[12] In the Appendix, Tables A.1 and A.2 present the list of all players and matches, respectively.

Table 2
Summary statistics.

Variables	Mean	Std. dev.	Min.	Max.	N
Panel A: match level					
Number of points	220.3	64.6	117	436	80
Number of games	36.4	10.1	20	77	80
Number of sets	3.6	0.7	3	5	80
Duration (in minutes)	150	48.9	69	284	80
Odds-implied winning probability difference	0.67	0.28	0.04	0.98	80
Difference in players' ATP ranking	41.9	43.9	1	223	80
Panel B: experiment level					
Number of points	45.8	14.7	17	97	320
Fraction of points to the receiver's left	0.55	0.12	0.15	0.87	320
Fraction of points to the receiver's right	0.44	0.12	0.12	0.85	320
Fraction of points won when played to the left	0.70	0.12	0.26	1	320
Fraction of points won when played to the right	0.70	0.14	0	1	320
Panel C: point level					
Number of shots per point	3.72	3.51	1	32	14,658
Serve speed (mph): first and second	111.0	13.02	54	143	14,484
Serve speed (mph): first	116.8	9.08	54	143	10,399
Serve speed (mph): second	96.0	8.87	56	133	4,085
Fraction of first serves	0.717	0.450	0	1	14,658

Notes: The table reports the summary statistics for the 80 men's singles matches in our sample. An experiment encompasses all the points played in each of the four distinct player-court combinations. For 174 points, IBM did not record the serve speed.

the right (70.1%). The result of a mean-comparison t-test confirms that the difference between these winning rates across choices is not statistically significant ($t = 0.301$, $p = 0.763$).

The difference in the odds-implied players' winning probabilities at the beginning of the match averages 67% and has a median of 78.6%. The minimum difference (4%) results from a quarterfinal match, whereas the maximum difference (98%) results from a first tournament stage match. Furthermore, the ranking difference for a match averages nearly 42 ranking points. Both the high average difference in winning probabilities and the high average ranking difference may be explained by the relatively high percentage (60%) of early tournament (pre-quarterfinal) matches in our sample, which are matches in which the chances of having a low-ranking player meeting a high-ranking player are higher. The individual player rankings vary between one and 236, with an average of 25 and a median of six (see Table A.1 for further details).

On average, an experiment consists of 45.8 points. In comparison, Walker and Wooders (2001) and Hsu et al. (2007) focus on lengthy matches, giving them a particularly long series of actions to analyze for each experiment: their experiments average

75.7 and 62.2 points, respectively. However, as both studies select matches from a period of over 20 years, their samples are affected by the important technological changes that have been adopted over the years that have altered how tennis is played. For example, new racket materials allow for faster shots, and new court materials such as polyester allow for faster rebounding and help to generate topspin (O'Donoghue, 2001). Cross and Pollard (2009) show that since 1998, the average men's serve speed significantly increased for all Grand Slam tournaments, finally converging to approximately 115 mph for all tournaments in 2006. In our sample, the average serve speed is 111 mph, the first serve being faster (117 mph) than the second (96 mph). On average, the percentage of first serves is roughly 72%.[13]

To make our results comparable with those of previous studies, we apply the methodology proposed by Walker and Wooders (2001) and described in Palacios-Huerta (2014, pp. 22–23). In a first step, we test the null hypothesis of the equality of a server's winning probability for each of the pure strategies available (left and right) by conducting the Pearson's chi-square goodness-of-fit test, at both the individual and joint levels, and the Kolmogorov-Smirnov (KS) test. Pearson's test is also known as a standard proportions test.

In the following equation, for each match, i will denote an experiment, which includes the series of actions chosen by a server from a certain court (left or right), i.e., from any player-court combination. In total, we analyze 320 experiments. We define p_i^j as the likelihood that a server from a given court will win a point when choosing the pure strategy $j \in \{L, R\}$, with n_i^j being the number of times that the server chooses j and $N_i^{j,W}$ and $N_i^{j,NW}$ being the number of times that the server wins the serve point (W) or not (NW), respectively, when choosing j. Under the null hypothesis, the server's probability of winning a point (or "winning rates") should be the same for the left serves as for the right serves ($H_0 : p_i^L = p_i^R = p_i$).

[13] If we do not drop the 472 double faults, the percentage of first serves decreases to 63.9%.

After replacing p_i with its maximum likelihood estimate $\frac{N_i^{L,W}+N_i^{R,W}}{n_i^L+n_i^R}$, we compute the following Pearson test statistic for experiment i:

$$P_i = \sum_{j\in\{L,R\}} \left[\frac{(N_i^{j,W} - n_i^j \cdot p_i)^2}{n_i^j \cdot p_i} + \frac{(N_i^{j,NW} - n_i^j \cdot (1-p_i))^2}{n_i^j \cdot (1-p_i)} \right], \tag{2}$$

which under the null hypothesis is distributed asymptotically as a χ^2 with one degree of freedom. Larger values of the Pearson statistic denote a tendency toward less equality in the winning proportions and, thus, a greater likelihood of rejecting the minimax prediction for that experiment.

Furthermore, we can also test the aggregate players' behavior, i.e., test whether each of the experiments is simultaneously generated by minimax play (Palacios-Huerta, 2014). Many rejections at the individual level do not necessarily imply a joint rejection. Importantly, the joint hypothesis that $p_i^L = p_i^R$ for *each* of the experiments allows for differences in probabilities p_i across experiments (Walker & Wooders, 2001). We compute the Pearson joint test statistic as the sum of the individual test statistics:

$$P = \sum_{i}^{320} P_i, \tag{3}$$

which under the null hypothesis is distributed asymptotically as a χ^2 with 320 degrees of freedom. Similar to the individual Pearson test, larger values of the chi-square statistic indicate a greater likelihood of rejecting the joint hypothesis.

Furthermore, Walker and Wooders (2001) note that under the joint null hypothesis that all observations were generated by minimax play, the $p-$values associated with the individual Pearson statistics (Eq. 2) should be 320 draws from the uniform distribution $U[0,1]$. Using the KS test statistic, we can test the null hypothesis that the hypothesized cumulative distribution function (CDF) of the Pearson $p-$values is uniform on [0,1] ($H_0 : F(x) = F_0(x) = x$). The test statistic is:

$$K = \sqrt{n} \cdot sup_{x \in [0,1]} |F_n(x) - x)|, \tag{4}$$

where $F_n(x)$ is the empirical CDF of the 320 $p-$values. As the KS test cannot be computed directly, we use the two-step procedure described in footnotes 17 and 18 of Walker and Wooders (2001, p. 1533). This test statistic has a known distribution (Mood, Graybill, & Boes, 1974, p. 509). Larger values of the KS test statistic (or smaller $p-$values) indicate a larger difference between the observed and theoretical distributions and, thus, a greater likelihood of rejecting the joint null hypothesis that each individual experiment is simultaneously generated by minimax play. Finally, the KS test can be visualized by juxtaposing all the 320 observed $p-$values associated with the Pearson statistics P_i with the uniform distribution. The larger the vertical distance between the two CDFs is, the larger the KS statistic and the more likely the rejection of the null hypothesis will be.

In a second step, we test the null hypothesis of the serial independence of the servers' choices, at both the individual and the joint levels, by implementing the runs test (Wald & Wolfowitz, 1940). We define a run as a succession of one or more identical actions (the serve direction) that are preceded and followed by a different choice or by no choice at all. For example, an experiment having the following sequence of actions $\{R, R, L, R, L, L\}$ has four runs. In a given service sequence, having too many (few) runs would suggest that the player is switching serve direction too often (not often enough), resulting in a series of choices that is negatively (positively) autocorrelated. In both cases, the test would suggest that the server is not changing direction in a random matter. A server who is following some sort of pattern could be exploited by his opponent, because the serve direction choices become predictable. A runs test would reject the null hypothesis of randomness if the number of runs is either too large or too small relative to the theory.

We assume a sequence of n elements of two actions, n_L for the serves to the left and n_R for the serves to the right, where $n_L + n_R = n$. Then, we compute for every player-court sequence of choices i the number of runs r_i. The probability distribution of the total amount of runs r in a random sample is known (Gibbons & Chakraborti, 2011). For large samples ($n > 20$), however, a good approximation to the null distribution is the normal distribution with mean $\mu_r = 1 + \frac{2n_L n_R}{n_L + n_R}$ and standard deviation $\sigma_r = \frac{2n_L n_R \cdot (2n_L n_R - n_L - n_R)}{(n_L + n_R)^2 \cdot (n_L + n_R - 1)}$. The limiting probability function of the test statistic $z = \frac{r - \mu_r}{\sigma_r}$ is the standard normal density. Using a continuity correction of 0.5, the left-tail and right-tail critical regions are as follows:

$$\frac{r + 0.5 - 1 - 2n_L n_R/n}{\sqrt{2n_L n_R(2n_L n_R - n)/[n^2(n-1)]}} \leq -z_\alpha, \tag{5}$$

and

$$\frac{r - 0.5 - 1 - 2n_L n_R/n}{\sqrt{2n_L n_R(2n_L n_R - n)/[n^2(n-1)]}} \geq z_\alpha, \tag{6}$$

where z_α is 0.025 (0.05) at the 5% (10%) significance level. If the test statistic is greater than its critical values, we reject the null hypothesis of randomness. Furthermore, we test the joint null hypothesis that all observations are serially independent in each of the 320 experiments using the KS test.

In general, when a strategic game is repeated over and over, some deviations from the equilibrium predictions are expected. To account for these deviations, we proceed as follows: if the percentage of experiments not consistent with *Prediction 1* or *2* exceeds 10% or, alternatively, 5% of all experiments, then we reject the null hypothesis of minimax.

5 Empirical results

5.1 Full sample results

***Prediction 1*: Testing for equality of winning probabilities**

In Table 3, columns (1) and (2) summarize the full sample results of the individual Pearson test.[14] Given the large number of experiments, we present the individual results in terms of the percentage of rejections at the standard 10% and 5% significance levels. Specifically, column (1) indicates the percentage of experiments in which the Pearson test $p-$value is smaller than 10%, whereas column (2) indicates the percentage of experiments in which the Pearson test $p-$value is smaller than 5%. As we previously explained, we reject *Prediction 1* when the actual percentage of the individual rejections is larger than 10% or, using a stricter cut-off, larger than 5%.

Table 3
Prediction 1: Full sample analysis.

			Pearson test: % of indiv. rej.		Joint statistics	
			(1)	(2)	(3)	(4)
Sample	Exp.	N	10%	5%	Pearson	KS
Full	320	14,658	12.81%	6.87%	407.3 (0.001)	1.769 (0.004)

Notes: The table reports the summary results of the Pearson and KS tests for *Prediction 1* using the full sample. An experiment (Exp.) consists of all the points played in any of the four server-court combinations. N indicates the total number of points used for testing. Columns (1) and (2) indicate the percentage of the individual events resulting in a rejection at the 10% and 5% levels, respectively. Columns (3) and (4) report the joint Pearson's chi-square goodness-of-fit and KS statistic ($p-$values in brackets), respectively.

For the full sample, the Pearson statistic is significant at the 10% level in 41 experiments and at the 5% level in 22 experiments, corresponding to 12.81% and 6.87% of the 320 experiments in total, respectively. The expected number of rejections according to the theory (i.e., when the minimax null hypothesis is true) is 36 at the 10% level and 18 at the 5% level. Thus, there are more rejections than expected under the null at the individual level.

At the aggregate level, we reject the joint null hypothesis of minimax play because both the joint Pearson and KS statistics, as reported in column (3) and column (4),

[14] In the Appendix A.2, Table A.3 presents the results for each experiment.

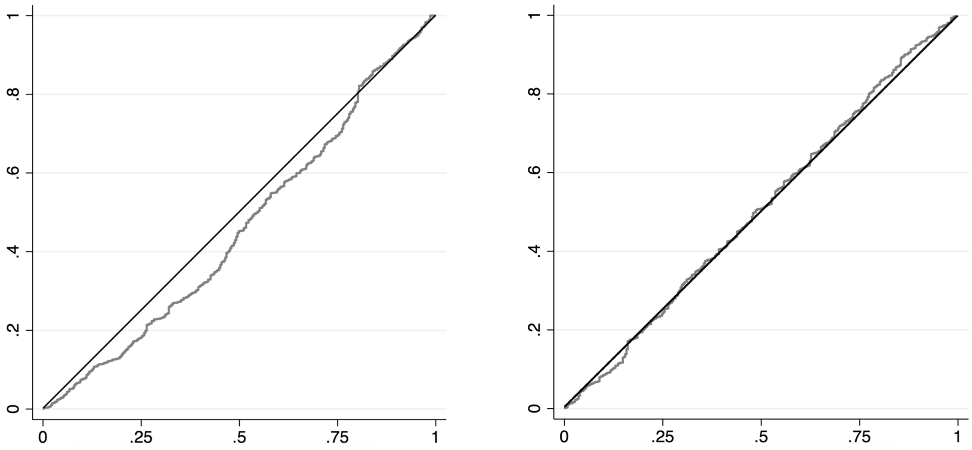


Figure 1
Kolmogorov-Smirnov tests for *Prediction 1* (left) and for *Prediction 2* (right). In both graphics, the closer the grey line is to the black 45° line, the closer the data are to the minimax predictions.

are large and have associated $p-$values less than 1%. Additionally, the graphic in the left part of Fig. 1 shows that the empirical CDF of the $p-$values associated with the 320 Pearson statistics P_i (represented by the grey line) does not precisely match the uniform c.d.f (represented by the black 45° line). Indeed, a larger vertical distance between the two lines indicates a larger KS statistic. Thus, we also reject the null hypothesis that each of the experiments is simultaneously generated by equilibrium play.

Overall, we do not find evidence supporting *Prediction 1.* This first result contradicts Hsu et al. (2007) and Walker and Wooders (2001), who find supporting evidence for *Prediction 1.*[15]

Prediction 2: Testing for serial independence

In Table 4, columns (1) through (4) summarize the results of the individual runs test, and column (5) shows the results of the joint runs KS test.[16] As in the previous section, we present the individual results in terms of the percentage of rejections

[15] As a comparison, in Walker and Wooders (2001), the percentage of rejections is 5% (at the 10% significance level) and 2.5% (at the 5% significance level), the joint Pearson statistic is 30.8 ($p-$value=0.852), and the KS statistic is 0.67 ($p-$value=0.76). In Hsu et al. (2007), for the men's singles sample, the percentage of rejections is 15% (at the 10% significance level) and 5% (at the 5% significance level), the joint Pearson statistic is 54.157 ($p-$value=0.067), and the KS statistic is 0.778 ($p-$value=0.580).

[16] In the Appendix A.2, Table A.4 presents the results for each experiment.

Table 4
Prediction 2: Full sample analysis.

			Runs test: % of individual rejections				Joint statistics
			(1)	(2)	(3)	(4)	(5)
Sample	Exp.	*N*	10%	+/–	5%	+/–	KS
Full	320	14,658	10.31%	17/16	4.68%	5/10	1.305 (0.072)

Notes: The table reports the summary results of the runs tests for the *Prediction 2* tests using the full sample. An experiment (Exp.) consists of all the points played of any of the four server-court combinations. *N* indicates the total number of points used for testing. The 10% and 5% columns indicate the number of rejections at the 10% and 5% significance levels, respectively. The "+/–" column indicates the number of experiments for which the null hypothesis is rejected either due to too many (+) or too few (–) runs.

at the standard 10% and 5% significance levels. Specifically, column (1) reports the percentage of experiments in which the null hypothesis of serial independence is rejected at the 10% significance level, whereas column (3) reports the percentage of rejections at the 5% significance level. Columns (2) and (4) report the number of individual rejections due to too many (+) or too few (–) runs at the 10% and 5% levels, respectively.

For the full sample, the individual data are very close to the second minimax prediction. The null hypothesis of serial independence is rejected at the 10% level for 33 experiments and at the 5% level for 15 experiments, corresponding to 10.31% and 4.68% of the 320 experiments in total, respectively. At the 10% significance level, we find one rejection more than expected under the null, whereas at the 5% significance level, we find one fewer than expected. According to column (4), 10 rejections out of 15 are caused by too few runs, i.e., when players do not switch often enough.

Concerning the question of whether the players' choices are jointly serially independent, the KS statistic is 1.305 and not statistically significant at the 5% level; thus, we cannot reject the null hypothesis that all observations are serially independent in each of the 320 experiments. This result is confirmed by the graphic in the right part of Fig. 1.

Overall, we find evidence supporting *Prediction 2*. This result is in line with Hsu et al. (2007) but not with Walker and Wooders (2001), who find that the server's choices are not random but negatively correlated.[17] Soccer studies, such as Palacios-Huerta (2003) and Dohmen and Sonnabend (2016), also find evidence supporting *Prediction 2*. However, penalty kicks are taken very infrequently, automatically resulting in a bias toward less serial correlation in the players' choices.

5.2 Subsample results

In this section, we retest *Predictions 1* and *2* with the same methodology but using various subsamples. Table 5 reports the subsample results for *Prediction 1.* For the lefty and righty matches subsamples, column (2) indicates that the percentage of rejections at the 5% level is almost twice as large for the lefty subsample (10.42%) than for the righty subsample (5.36%). However, the joint Pearson and KS statistics are larger for the righty matches. In general, these results confirm the idea that left-handedness is an important physical characteristic that has to be taken into consideration when analyzing players' strategic choices in tennis.

In both the early and late tournament stage subsamples, the percentage of individual rejections is higher than predicted by the minimax theorem. The joint Pearson and KS statistics are significant for both subsamples, leading to a rejection of the null hypothesis of minimax. Importantly, when comparing the two groups, we note that the players play closer to the minimax predictions in the early tournament stages. This result contrasts with that of Walker and Wooders (2001), who analyze mostly finals and find evidence supporting *Prediction 1.*

When we compare the early- and late-set subsamples, both the individual and joint Pearson tests show that players behave very close to the minimax predictions in the later sets.[18] This result is valuable because it suggests that players play closer

[17] As a comparison, in Walker and Wooders (2001), the percentage of rejections is 20% (at the 10% significance level) and 12.5% (at the 5% significance level), and the joint KS statistic is 1.948 ($p-$value=0.001). In Hsu et al. (2007), for the men's singles sample, the percentage of rejections is 10% (at the 10% significance level) and 5% (at the 5% significance level), and the joint KS statistic is 0.778 (5% critical value=1.328).

[18] All Grand Slam matches are played as the best of five sets. However, unlike the fourth and fifth sets, the third set can end a match only when a player leads 2–0 after the second set. Thus, in a further test, we include the third set in the first group only if one player had a 2–0 lead after the second set. The results are similar.

Table 5
Prediction 1: Subsample analyses.

			Pearson test: % of indiv. rej.		Joint statistics	
			(1)	(2)	(3)	(4)
Sample	Exp.	N	10%	5%	Pearson	KS
Lefty	96	4,139	16.67%	10.42%	133.93 (0.006)	1.18 (0.119)
Righty	224	10,519	11.16%	5.36%	273.35 (0.013)	1.49 (0.023)
Early stages	248	10,872	12.10%	6.05%	299.45 (0.014)	1.53 (0.018)
Late stages	72	3,876	15.28%	9.72%	107.83 (0.003)	1.29 (0.069)
Early sets	320	9,502	15.62%	8.12%	399.95 (0.001)	1.93 (0.001)
Late sets	316	5,156	10.13%	4.11%	327.58 (0.314)	1.41 (0.043)
Short matches	160	5,575	10.00%	6.25%	194.48 (0.032)	1.61 (0.010)
Long matches	160	9,083	15.62%	7.50%	212.80 (0.003)	1.03 (0.234)
Balanced	80	3,806	15.00%	11.25%	130.80 (0.000)	1.53 (0.017)
Unbalanced	80	3,274	5.00%	5.00%	80.88 (0.451)	0.74 (0.634)

Notes: The table reports the summary results of the Pearson and KS tests for *Prediction 1* using various subsamples. An experiment (Exp.) consists of all the points played in any of the four server-court combinations. N indicates the total number of points used for testing. Columns (1) and (2) indicate the percentage of the individual events resulting in a rejection at the 10% and 5% level, respectively. Columns (3) and (4) report the joint Pearson's chi-square goodness-of-fit and KS statistic ($p-$values in brackets), respectively.

to minimax at later and more important moments of the match, possibly indicating that they cope well with pressure and that they improve their strategic skills over the match.

Concerning the short and long match subsamples, column (1) indicates that the percentage of rejections at the 10% level is 15.62% in long matches and 10% in short matches. At the 5% level, the difference between the two groups is small. In contrast to the joint Pearson test, the joint KS test does not lead to a rejection of minimax for the long matches. Overall, it appears that the players play closer to *Prediction 1* in short matches than in long matches. This result may be attributed to fatigue, which influences the players' ability to play as intended (Davey et al., 2002).

Perhaps more surprisingly, we find that the percentage of rejections at the 10% level is three times larger for the balanced matches than for the unbalanced matches. Column (2) confirms this tendency as well at the 5% level. At the aggregate level, we reject the joint null hypothesis of minimax for the balanced group at the 1% level per both the joint Pearson and the KS tests. This finding is particularly interesting not only because of the large difference in the percentage of rejections between the 25% most balanced and the 25% most unbalanced matches but also because of the closeness of the results to the first minimax prediction, both at the individual and joint level, for the unbalanced matches. We further analyze this interesting finding at the end of this section.

Overall, we observe a tendency toward accepting *Prediction 1*, which states that each player's probability of winning a point on serve across choices should be the same in matches between right-handed players, in the early stages of a tournament but in late sets of a match, and in shorter and unbalanced matches.

Table 6 reports the subsample results for *Prediction 2*. For the lefty and righty matches subsamples, columns (1) and (3) indicate that *Prediction 2* is rejected at the individual level only for the matches with a left-handed player. In the righty matches, the percentage of individual rejections is 9.28% at the 10% significance level and 4.46% at the 5% significance level. The results at the experiment level for the lefty matches are slightly higher. However, at the joint level, we reject the joint null hypothesis only for the righty matches.

At the 5% level, the percentage of rejections of the runs test in the early-stage matches (5.24%) is almost double that in the late-stage matches. Taken together, our results suggest that whereas the players play according to *Prediction 1* in the early stages, they play according to *Prediction 2* in the later stages.

In contrast to the early- and late-stage subsamples, we find that both *Predictions 1* and *2* fare better in later sets than in earlier sets. Thus, by showing that both minimax predictions describe the players' behavior better in the later moments of a match, our findings provide evidence confirming Binmore et al.'s (2001) suggestion that people come closer to equilibrium by undergoing a trial-and-error adjustment process.

Fatigue also appears to play a role in the players' ability to randomize their serves.

Table 6
Prediction 2: Subsample analyses.

			Runs test: % of individual rejections				Joint statistics
			(1)	(2)	(3)	(4)	(5)
Sample	Exp.	*N*	10%	+/–	5%	+/–	KS
Lefty	96	4,139	11.45%	8/3	5.20%	2/3	1.10 (0.180)
Righty	224	10,519	9.82%	9/13	4.46%	3/7	1.88 (0.002)
Early stages	248	10,872	10.80%	16/11	5.24%	5/8	1.08 (0.202)
Late stages	72	3,786	8.30%	1/5	2.70%	0/2	1.48 (0.026)
Early sets	320	9,502	6.56%	10/11	4.06%	7/6	1.00 (0.273)
Late sets	316	5,156	5.69%	10/8	2.84%	6/3	0.71 (0.685)
Short matches	160	5,575	6.87%	7/4	2.50%	1/3	0.97 (0.309)
Long matches	160	9,083	13.75%	10/12	6.87%	4/7	1.34 (0.059)
Balanced	80	3,806	13.75%	5/6	6.25%	2/3	0.87 (0.437)
Unbalanced	80	3,274	5.00%	3/1	2.50%	1/1	0.67 (0.739)

Notes: The table reports the summary results of the runs tests for *Prediction 2* using various subsamples. An experiment (Exp.) consists of all the points played in any of the four server-court combinations. *N* indicates the total number of points used for testing. The 10% and 5% columns indicate the number of rejections at the 10% and 5% significance levels, respectively. The "+/–" column indicates the number of experiments for which the null hypothesis is rejected either due to too many (+) or too few (–) runs.

Concerning the short- and long-match subsamples, columns (1) and (3) indicate that the percentage of rejections both at the 10% level and at the 5% level is more than twice as large in long matches. Considering the joint statistics, the KS test leads to a rejection of the joint null hypothesis only for the long matches.

For the balanced and unbalanced matches subsamples, the percentage of individual rejections in the balanced group (6.25%) is twice as large as in the unbalanced group (2.50%). The difference in the percentage of individual rejections between subsamples is even larger at the 10% significance level (8.75%). Although we cannot reject the null hypothesis in any of the groups according to the joint KS test, the KS statistic is smaller for the unbalanced group. The players' behavior is very different between these two groups for both the first and the second minimax prediction. This finding confirms the intuition that playing against an equally skilled opponent is not the same as playing against a less or more skilled opponent.

Overall, we observe a tendency toward accepting *Prediction 2*, which states that the server's choices are serially independent in matches between right-handed players, in the late stages of a tournament, in the later sets of a match, and in shorter and unbalanced matches. When jointly considering the results for *Predictions 1* and *2*, the tournament stage subsample is the only subsample that provides contradictory results concerning minimax play—we find a tendency of playing according to *Prediction 1* in the early stages but a tendency of playing according to *Prediction 2* in the late stages. Thus, to summarize our subsample results, we find a tendency toward minimax play in matches between right-handed players, in the later sets of a match, in shorter, and in unbalanced matches.

Since our previous subsample definition of balanced and unbalanced matches does not account for the relative skill level of each player—for example, a match between two top players is a balanced match, as is one between two non-top players—we now define four further subsamples depending on the skills of the server and of the receiver: top server vs. top receiver, top server vs. non-top receiver, non-top server vs. non-top receiver, and non-top server vs. top receiver. For each of these four subsamples, Panel A in Table 7 reports the results for *Prediction 1*, and Panel B reports the results for *Prediction 2*. In general, the percentage of rejections varies considerably across the four subsamples, confirming the importance of considering the skills of both players.

In particular, the "top server vs. non-top receiver" subsample is the only one for which *Predictions 1* and *2* are consistently rejected, both at the individual and joint levels. In other words, it appears that top servers do not play near the predictions of minimax when facing non-top receivers. As this looks surprising, we further analyze this finding in the following subsection.

5.3 Winning likelihood, minimax, and the players' skills

To win, it might be rational for a player to deviate when his opponent systematically deviates or is expected to deviate from minimax play (Levitt et al., 2010; Palacios-Huerta, 2014). If this is true, we should find that playing minimax in the 124 "top

Table 7
Predictions 1 and *2*: Relative skills subsample analyses.

Panel A: *Prediction 1*

			Pearson test: % of indiv. rej.		Joint statistics	
			(1)	(2)	(3)	(4)
Sample	Exp.	N	10%	5%	Pearson	KS
Top server vs. top receiver	32	1,699	21.87%	15.62%	61.17 (0.001)	1.19 (0.114)
Top server vs. non-top receiver	124	5,415	11.29%	6.45%	159.29 (0.017)	1.45 (0.029)
Non-top server vs. non-top receiver	44	1,957	15.90%	9.09%	51.59 (0.201)	0.54 (0.923)
Non-top server vs. top receiver	120	5,587	10.83%	4.16%	135.24 (0.161)	0.94 (0.331)

Panel B: *Prediction 2*

			Runs test: % of individual rejections				Joint statistics
			(5)	(6)	(7)	(8)	(9)
Sample	Exp.	N	10%	+/−	5%	+/−	KS
Top server vs. top receiver	32	1,699	9.37%	1/2	6.25%	1/1	1.24 (0.100)
Top server vs. non-top receiver	124	5,415	14.51%	7/11	5.64%	1/6	2.43 (0.000)
Non-top server vs. non-top receiver	44	1,957	6.81%	3/0	2.27%	1/0	0.85 (0.462)
Non-top server vs. top receiver	120	5,587	7.50%	6/3	4.16%	2/3	1.37 (0.051)

Notes: Panel A reports the summary results of the Pearson and KS tests for *Prediction 1*. Panel B reports the summary results of the runs tests for *Prediction 2*. An experiment (Exp.) consists of all the points (N) played by a given server against a given server from a given court.

server vs. non-top receiver" experiments actually *reduces* the likelihood of winning a match for the top player. We analyze this hypothesis in two steps.

We begin by asking whether playing minimax affects, on average, a player's likelihood of winning a match. Formally, we estimate the following regression:

$$win\_match_i = \alpha_0 + \beta_1 \cdot minimax_i + \beta_2 \cdot control_i + u_i \ , \tag{7}$$

where i indexes an experiment. The variable win_match equals one when the server wins the match and is zero otherwise. For each experiment i, we know whether the individual Pearson and runs tests result in a rejection of *Prediction 1* or *2* at the 10% level. If both *Prediction 1* and *2* are satisfied for experiment i, then the $minimax$ variable equals one and is zero otherwise. To control for other factors influencing the probability of the server winning the match, we also include the variable $control$, corresponding to the odds-implied winning probability of the server at the beginning of the match. This is appropriate because the betting odds reflect all public information available about the two players, such as surface preferences, recent injuries, or current winning streak (Wolfers & Zitzewitz, 2006).

We then ask whether it is rational for top servers to play minimax against non-top receivers, given their goal of winning the match. We interact the *minimax* variable with the *topbad* dummy, which equals one for all the "top server vs. non-top receiver" experiments. Thus, we estimate the following regression:

$$\begin{aligned} win\_match_i = \alpha_0 + \beta_1 \cdot minimax_i + \beta_2 \cdot topbad_i \\ + \beta_3 \cdot (minimax \cdot topbad_i) + \beta_4 \cdot control_i + u_i \ . \end{aligned} \tag{8}$$

The first two columns of Table 8 report the ordinary least squares regression estimates for both regression equations.[19] Column (1) indicates that on average, playing minimax does not have a significant effect on the likelihood of winning. However, column (2) indicates that a top server decreases his winning likelihood by more than 18% if he plays minimax against non-top receivers. In other words, this confirms the idea that more deviations from minimax play do not hurt top servers, possibly because either top players have more sophisticated tennis skills that compensate for suboptimal minimax behavior or non-top players do not have the required skills to readily exploit that suboptimal behavior.

As a further test, we replace the dependent variable *win_match* in Eq. 7 and 8 with the total number of points won in each experiment (*points_won*) by the server. The advantage of using *points_won* is that it contains more information than *win_match* and varies at the experiment level. Column (4) in Table 8 shows that top players lose, on average, more than six points when playing minimax against non-top players.

A limitation of this analysis is that we regress the outcome of the match for the

[19] As a robustness check, we use logistic regression. The results are similar.

Table 8
Minimax play and winning.

	Dependent variable: *win_match*		Dependent variable: *points_won*	
	(1)	(2)	(3)	(4)
minimax	-0.008	0.067	-1.714	0.762
	(0.038)	(0.052)	(1.539)	(1.917)
topbad		0.068		-0.340
		(0.110)		(3.727)
minimax $\times$ *topbad*		-0.186**		-6.229**
		(0.075)		(3.162)
control	1.020***	1.105***	53.668***	59.226***
	(0.053)	(0.106)	(2.604)	(3.951)
Exp.	320	320	320	320
N	14,658	14,658	14,658	14,658
F-stat	188.55	111.71	216.08	116.43
$Pr > F$	0.000	0.000	0.000	0.000
R^2	0.560	0.568	0.708	0.714

Notes: The table reports the results of cross-sectional OLS regressions. The dependent variable in columns (1) and (2) is a dummy indicating whether the server won the match, and that in columns (3) and (4) is a variable indicating the total number of points won by the server from a given court. All estimations also include a constant (not reported). The data are at the experiment level (Exp.=320) and include 80 men's singles matches played at Wimbledon. Robust standard errors that have been adjusted for clustering at the match, and player levels are provided in parentheses. In all models, *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

server on a behavior that is not determined solely by the server (*minimax*). In particular, when testing *Prediction 1*, we cannot directly observe whether the receiver is playing to make the server's payoffs equal across choices. Concerning *Prediction 2*, where we analyze the randomizing skills of the servers, our analysis should be less problematic. Although the receiver can attempt to influence the server's behavior by moving across the court to incentivize the server to hit the ball in a given direction, we believe that the decision to play according to *Prediction 2* is more in the hands of the server.

6 Concluding remarks

Earlier field studies have not yet explained *when* professionals behave according to minimax and the role of the players' skills level. Using our full sample, we find that while players' choices are serially independent (*Prediction 2*), the winning probabilities are not statistically identical across choices (*Prediction 1*). Importantly, we show that players tend toward minimax play in matches between right-handed players, in the later sets of a match, and in shorter and unbalanced matches. Here, further research is needed to clarify how the specific conditions of tennis are generalizable: for example, how do our results for the "later sets of a match" subsample apply to general situations under pressure?

After further investigation of the importance of the skills of the players, we find that situations involving a top server playing against a non-top receiver persistently result in more deviations from minimax play. Crucially, our results show that top servers can significantly increase their likelihood of winning the match and the number of points won by, at times, deviating from the minimax equilibrium. Thus, future laboratory studies should take into consideration the experiment participants' skills when analyzing their behavior in strategic games.

A Appendix

A.1 Tennis

The International Tennis Federation (ITF) describes the tennis serving rules. At Grand Slam tournaments, matches are played as the best of five sets: the first player to win three sets wins the match. At the beginning of the match, a coin toss decides which player starts serving in the first game. In the second game, his opponent will serve and so forth. The player who loses a set will start serving in the next set.

When serving in a standard game, the server shall stand behind alternate halves of the court, starting from the right half of the court in every game. The serve shall pass over the net and hit the service court diagonally opposite, i.e., the left (*ad*) or the right (*deuce*) court, before the receiver returns it. If the first serve is a fault, the server shall serve again (the second serve) without delay from behind the same half of the court from which that fault was served.

Fig. A.1 shows the tennis court, which for singles matches is 78 feet (23.77 meters) long and 27 feet (8.23 meters) wide, and illustrates possible serve directions (excluding the serve to the center). The server shall stand at rest with both feet behind the baseline (the black point approximately indicates his position) and put the ball into the cross-court service box (the grey area). The receiver is free to choose his position. The court on the left shows the server's two choices for right (*deuce*) points, that is, serving close to the center line (L) toward the receiver's left or serving wide toward the receiver's right (R). The court on the right shows the server's two choices for left (*ad*) points, that is, serving close to the center line toward the receiver's right (R) or serving wide toward the receiver's left (L).

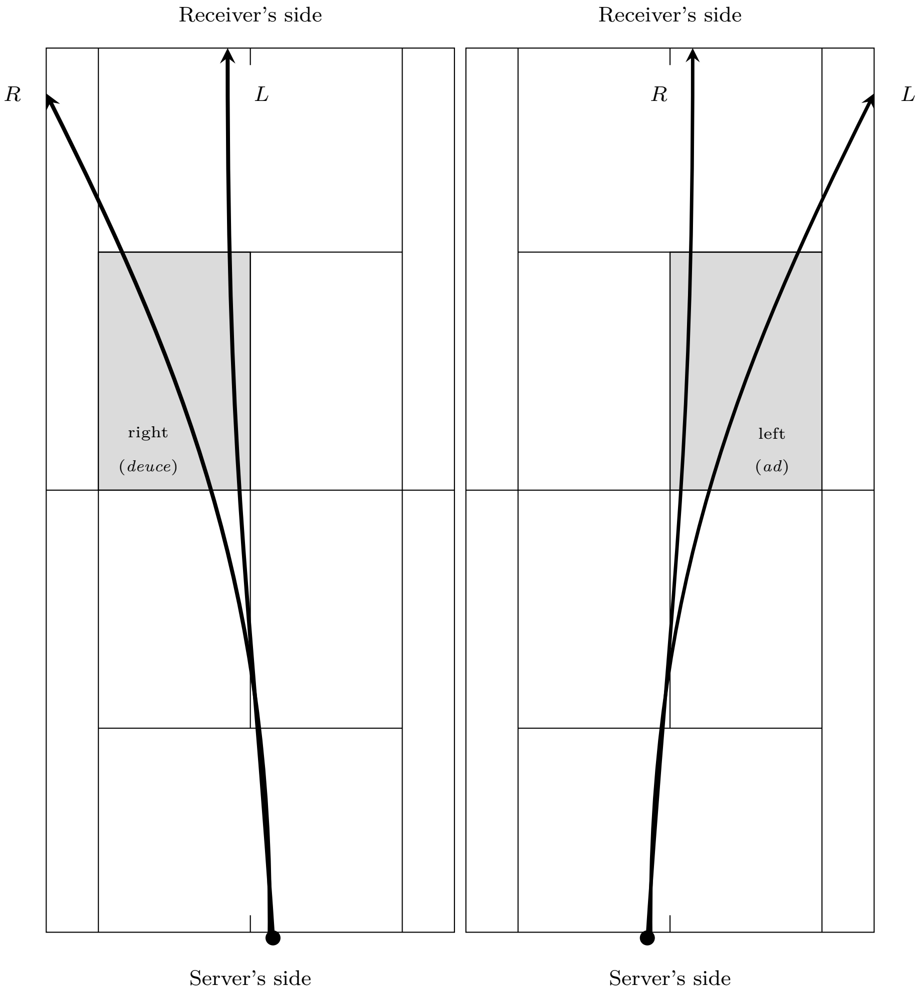


Figure A.1
The tennis court and illustrative serve choices.

Table A.1
Players' descriptive statistics.

Player	Left-handed	Matches	Ranking	Age	Player	Left-handed	Matches	Ranking	Age
Martin Alund	0	1	101.0	27.0	Jesse Levine	1	2	122.5	23.0
Marcos Baghdatis	0	1	30.0	26.0	Michael Llodra	1	1	35.0	31.0
Julien Benneteau	0	2	32.5	30.5	Paolo Lorenzi	0	1	83.0	32.0
Tomas Berdych	0	4	11.3	25.5	Yen-Hsun Lu	0	1	47.0	30.0
Ilija Bozoljac	0	1	152.0	24.0	Xavier Malisse	0	1	75.0	31.0
Arnaud Clement	0	1	86.0	32.0	Adrian Mannarino	1	1	55.0	23.0
Steve Darcis	0	1	135.0	29.0	Florian Mayer	0	1	34.0	29.0
Juan Martin Del Potro	0	3	8.0	24.0	Jurgen Melzer	1	3	30.0	30.5
Grigor Dimitrov	0	4	12.8	23.0	Gilles Muller	1	1	103.0	31.0
Novak Djokovic	0	16	1.9	24.5	Andy Murray	0	16	3.7	24.5
Alejandro Falla	1	1	88.0	26.0	Rafael Nadal	1	10	1.4	26.0
Roger Federer	0	30	2.8	29.5	David Nalbandian	0	1	23.0	29.0
David Ferrer	0	2	4.5	30.5	Philipp Petzschner	0	1	41.0	26.0
Mardy Fish	0	1	9.0	29.0	Albert Ramos	1	2	55.0	24.5
Fabio Fognini	0	1	68.0	25.0	Milos Raonic	0	1	9.0	23.0
Guillermo Garcia Lopez	0	1	42.0	26.0	Bobby Reynolds	0	1	156.0	30.0
Richard Gasquet	0	2	11.0	26.0	Andy Roddick	0	3	6.0	26.0
Santiago Giraldo	0	1	35.0	26.0	Blaz Rola	1	1	92.0	23.0
David Goffin	0	1	105.0	23.0	Lukas Rosol	0	1	52.0	28.0
Ernests Gulbis	0	1	39.0	24.0	Luke Saville	0	1	236.0	20.0
Tommy Haas	0	2	34.0	31.0	Gilles Simon	0	1	44.0	29.0
Victor Hanescu	0	1	48.0	31.0	Robin Soderling	0	2	9.0	24.5
Ryan Harrison	0	1	150.0	22.0	Go Soeda	0	1	129.0	28.0
Lleyton Hewitt	0	2	41.0	28.5	Joao Sousa	0	1	41.0	25.0
Denis Istomin	0	1	45.0	27.0	Sergiy Stakhovsky	0	1	116.0	27.0
Jerzy Janowicz	0	2	22.0	22.0	Jo Wilfried Tsonga	0	5	14.2	26.0
Ivo Karlovic	0	1	36.0	30.0	Fernando Verdasco	1	2	54.0	29.0
Philipp Kohlschreiber	0	1	32.0	25.0	Stan Wawrinka	0	6	8.2	26.5
Mikhail Kukushkin	0	2	62.0	24.5	Mikhail Youzhny	0	3	25.3	30.0

Notes: The table reports the players' descriptive statistics (sorted alphabetically by last name). We indicate whether the players are left-handed (1=yes), in how many matches they played, their ATP ranking (or the average ATP ranking if they played at Wimbledon in more than one year), and age (or the average age if they played at Wimbledon in more than one year). In total, our sample contains 58 unique players.

Table A.2
List of the matches included in our sample.

Player 1	Player 2	Date	Stage	Player 1	Player 2	Date	Stage
G.Garcia-Lopez	R.Federer	24-Jun-2009	2	D.Ferrer	A.Murray	4-Jul-2012	5
P.Kohlschreiber	R.Federer	26-Jun-2009	3	A.Murray	J.Tsonga	6-Jul-2012	6
S.Wawrinka	J.Levine	27-Jun-2009	3	N.Djokovic	R.Federer	6-Jul-2012	6
A.Murray	S.Wawrinka	29-Jun-2009	4	R.Federer	A.Murray	8-Jul-2012	7
R.Soderling	R.Federer	29-Jun-2009	4	V.Hanescu	R.Federer	24-Jun-2013	1
T.Haas	N.Djokovic	1-Jul-2009	5	R.Nadal	S.Darcis	24-Jun-2013	1
L.Hewitt	A.Roddick	1-Jul-2009	5	N.Djokovic	F.Mayer	25-Jun-2013	1
I.Karlovic	R.Federer	1-Jul-2009	5	D.Ferrer	M.Alund	25-Jun-2013	1
T.Haas	R.Federer	3-Jul-2009	6	A.Ramos	J.Del Potro	25-Jun-2013	1
A.Roddick	A.Murray	3-Jul-2009	6	J-W.Tsonga	E.Gulbis	26-Jun-2013	2
A.Roddick	R.Federer	5-Jul-2009	7	F.Verdasco	J.Benneteau	26-Jun-2013	2
R.Federer	A.Falla	21-Jun-2010	1	N.Djokovic	B.Reynolds	27-Jun-2013	2
R.Federer	I.Bozoljac	23-Jun-2010	2	R.Gasquet	G.Soeda	27-Jun-2013	2
R.Federer	A.Clement	25-Jun-2010	3	J.Levine	J.Del Potro	27-Jun-2013	2
P.Petzschner	R.Nadal	26-Jun-2010	3	J.Melzer	S.Stakhovsky	28-Jun-2013	3
N.Djokovic	L.Hewitt	28-Jun-2010	4	J.Janowicz	J.Melzer	1-Jul-2013	4
R.Federer	J.Melzer	28-Jun-2010	4	M.Youzhny	A.Murray	1-Jul-2013	4
R.Federer	T.Berdych	30-Jun-2010	5	N.Djokovic	T.Berdych	3-Jul-2013	5
R.Soderling	R.Nadal	30-Jun-2010	5	F.Verdasco	A.Murray	3-Jul-2013	5
J.Tsonga	A.Murray	30-Jun-2010	5	N.Djokovic	J.Del Potro	5-Jul-2013	6
A.Murray	R.Nadal	2-Jul-2010	6	J.Janowicz	A.Murray	5-Jul-2013	6
T.Berdych	N.Djokovic	2-Jul-2010	6	N.Djokovic	A.Murray	7-Jul-2013	7
T.Berdych	R.Nadal	4-Jul-2010	7	G.Dimitrov	R.Harrison	23-Jun-2014	1
M.Kukushkin	R.Federer	21-Jun-2011	1	A.Murray	D.Goffin	23-Jun-2014	1
A.Mannarino	R.Federer	23-Jun-2011	2	P.Lorenzi	R.Federer	24-Jun-2014	1
M.Baghdatis	N.Djokovic	25-Jun-2011	3	S.Wawrinka	J.Sousa	24-Jun-2014	1
D.Nalbandian	R.Federer	25-Jun-2011	3	G.Dimitrov	L.Saville	25-Jun-2014	2
A.Murray	R.Gasquet	27-Jun-2011	4	A.Murray	B.Rola	25-Jun-2014	2
M.Llodra	N.Djokovic	27-Jun-2011	4	G.Muller	R.Federer	26-Jun-2014	2
M.Youzhny	R.Federer	27-Jun-2011	4	L.Rosol	R.Nadal	26-Jun-2014	2
R.Nadal	M.Fish	29-Jun-2011	5	S.Wawrinka	Y-H.Lu	26-Jun-2014	2
J.Tsonga	R.Federer	29-Jun-2011	5	N.Djokovic	G.Simon	27-Jun-2014	3
R.Nadal	A.Murray	1-Jul-2011	6	S.Giraldo	R.Federer	28-Jun-2014	3
J.Tsonga	N.Djokovic	1-Jul-2011	6	M.Kukushkin	R.Nadal	28-Jun-2014	3
R.Nadal	N.Djokovic	3-Jul-2011	7	S.Wawrinka	D.Istomin	30-Jun-2014	3
R.Federer	A.Ramos	25-Jun-2012	1	A.Murray	G.Dimitrov	2-Jul-2014	5
R.Federer	F.Fognini	27-Jun-2012	2	S.Wawrinka	R.Federer	2-Jul-2014	5
R.Federer	J.Benneteau	29-Jun-2012	3	N.Djokovic	G.Dimitrov	4-Jul-2014	6
R.Federer	X.Malisse	2-Jul-2012	4	R.Federer	M.Raonic	4-Jul-2014	6
R.Federer	M.Youzhny	4-Jul-2012	5	N.Djokovic	R.Federer	6-Jul-2014	7

Notes: The table lists all 80 matches in the sample played at Wimbledon. Stage indicates the tournament stage, where “1” indicates a first round match, “7” a final, and so on.

A.2 Full sample results

Tables A.3 and A.4 present the the full sample results (** indicates a rejection at the 10% level, * indicates a rejection at the 5% level). In Table A.3, "Serves" indicates the total number of serve points played by a server from a given court; "Mixture" is computed by dividing the number of serves to the left or right by the total number of serves; "Won" indicates the total number of serves won, either to the left or right; "Win rates", the key variable for testing *Prediction 1*, is computed by dividing "Won" by "Serves" (for left or right). For example, Garcia-Lopez served 15 times to the left from the ad-court and won ten points, yielding a win rate of roughly 67% from that court. Finally, the Pearson statistic is computed according to Eq. 2.

Table A.3
Prediction 1: Full sample results.

			Serves			Mixture		Won		Win rates		Individual test	
Year	Server	Court	L	R	Total	L	R	L	R	L	R	Pearson	p-value
24-Jun-09	G.Garcia-Lopez	Ad	15	11	26	0.58	0.42	10	4	0.67	0.36	2.345	0.125
24-Jun-09	G.Garcia-Lopez	Deuce	20	11	31	0.65	0.35	13	10	0.65	0.91	2.488	0.114
24-Jun-09	R.Federer	Ad	20	13	33	0.61	0.39	14	12	0.70	0.92	2.346	0.125
24-Jun-09	R.Federer	Deuce	21	13	34	0.62	0.38	15	12	0.71	0.92	2.141	0.143
26-Jun-09	P.Kohlschreiber	Ad	27	22	49	0.55	0.45	14	11	0.52	0.50	0.017	0.897
26-Jun-09	P.Kohlschreiber	Deuce	23	25	48	0.48	0.52	13	14	0.57	0.56	0.001	0.970
26-Jun-09	R.Federer	Ad	30	25	55	0.55	0.45	20	16	0.67	0.64	0.043	0.835
26-Jun-09	R.Federer	Deuce	31	27	58	0.53	0.47	23	21	0.74	0.78	0.101	0.750
27-Jun-09	S.Wawrinka	Ad	24	29	53	0.45	0.55	20	16	0.83	0.55	4.780	0.028*
27-Jun-09	S.Wawrinka	Deuce	16	38	54	0.30	0.70	12	26	0.75	0.68	0.234	0.628
27-Jun-09	J.Levine	Ad	37	27	64	0.58	0.42	26	17	0.70	0.63	0.378	0.538
27-Jun-09	J.Levine	Deuce	24	45	69	0.35	0.65	15	27	0.63	0.60	0.041	0.839
29-Jun-09	A.Murray	Ad	39	22	61	0.64	0.36	27	19	0.69	0.86	2.227	0.135
29-Jun-09	A.Murray	Deuce	29	37	66	0.44	0.56	17	25	0.59	0.68	0.562	0.453
29-Jun-09	S.Wawrinka	Ad	38	34	72	0.53	0.47	22	18	0.58	0.53	0.178	0.672
29-Jun-09	S.Wawrinka	Deuce	38	38	76	0.50	0.50	28	25	0.74	0.66	0.561	0.453
29-Jun-09	R.Soderling	Ad	22	17	39	0.56	0.44	19	11	0.86	0.65	2.534	0.111
29-Jun-09	R.Soderling	Deuce	14	20	34	0.41	0.59	12	19	0.86	0.95	0.883	0.347
29-Jun-09	R.Federer	Ad	24	20	44	0.55	0.45	22	16	0.92	0.80	1.261	0.261
29-Jun-09	R.Federer	Deuce	22	24	46	0.48	0.52	15	19	0.68	0.79	0.718	0.396
01-Jul-09	T.Haas	Ad	20	25	45	0.44	0.56	16	18	0.80	0.72	0.385	0.534
01-Jul-09	T.Haas	Deuce	33	21	54	0.61	0.39	26	17	0.79	0.81	0.037	0.847
01-Jul-09	N.Djokovic	Ad	28	18	46	0.61	0.39	21	14	0.75	0.78	0.046	0.829
01-Jul-09	N.Djokovic	Deuce	25	24	49	0.51	0.49	22	15	0.88	0.63	4.306	0.037*
01-Jul-09	L.Hewitt	Ad	37	42	79	0.47	0.53	21	26	0.57	0.62	0.216	0.641
01-Jul-09	L.Hewitt	Deuce	25	48	73	0.34	0.66	23	37	0.92	0.77	2.499	0.113
01-Jul-09	A.Roddick	Ad	44	33	77	0.57	0.43	31	23	0.70	0.70	0.005	0.942
01-Jul-09	A.Roddick	Deuce	49	25	74	0.66	0.34	40	17	0.82	0.68	1.739	0.187
01-Jul-09	I.Karlovic	Ad	19	15	34	0.56	0.44	15	12	0.79	0.80	0.006	0.939
01-Jul-09	I.Karlovic	Deuce	22	16	38	0.58	0.42	16	14	0.73	0.88	1.216	0.270
01-Jul-09	R.Federer	Ad	20	15	35	0.57	0.43	17	13	0.85	0.87	0.019	0.889
01-Jul-09	R.Federer	Deuce	28	12	40	0.70	0.30	26	12	0.93	1.00	0.902	0.342
03-Jul-09	T.Haas	Ad	27	17	44	0.61	0.39	19	11	0.70	0.65	0.154	0.694
03-Jul-09	T.Haas	Deuce	37	12	49	0.76	0.24	24	9	0.65	0.75	0.423	0.515
03-Jul-09	R.Federer	Ad	25	14	39	0.64	0.36	22	13	0.88	0.93	0.230	0.631
03-Jul-09	R.Federer	Deuce	20	18	38	0.53	0.47	19	14	0.95	0.78	2.459	0.116
03-Jul-09	A.Roddick	Ad	29	32	61	0.48	0.52	19	25	0.66	0.78	1.203	0.272
03-Jul-09	A.Roddick	Deuce	37	25	62	0.60	0.40	29	15	0.78	0.60	2.446	0.117
03-Jul-09	A.Murray	Ad	45	18	63	0.71	0.29	34	10	0.76	0.56	2.442	0.118
03-Jul-09	A.Murray	Deuce	45	17	62	0.73	0.27	33	14	0.73	0.82	0.547	0.459
05-Jul-09	A.Roddick	Ad	41	53	94	0.44	0.56	28	41	0.68	0.77	0.973	0.323
05-Jul-09	A.Roddick	Deuce	73	24	97	0.75	0.25	62	15	0.85	0.63	5.553	0.018*
05-Jul-09	R.Federer	Ad	47	33	80	0.59	0.41	38	27	0.81	0.82	0.012	0.913
05-Jul-09	R.Federer	Deuce	45	42	87	0.52	0.48	36	36	0.80	0.86	0.497	0.480

Continued on next page

Table A.3 – *Continued from previous page*

			Serves			Mixture		Won		Win rates		Individual test	
Year	Server	Court	L	R	Total	L	R	L	R	L	R	Pearson	p-value
21-Jun-10	R.Federer	Ad	28	39	67	0.42	0.58	18	29	0.64	0.74	0.790	0.374
21-Jun-10	R.Federer	Deuce	30	39	69	0.43	0.57	20	31	0.67	0.79	1.445	0.229
21-Jun-10	A.Falla	Ad	71	12	83	0.86	0.14	42	9	0.59	0.75	1.088	0.296
21-Jun-10	A.Falla	Deuce	72	10	82	0.88	0.12	45	5	0.63	0.50	0.577	0.447
23-Jun-10	R.Federer	Ad	31	26	57	0.54	0.46	24	19	0.77	0.73	0.144	0.704
23-Jun-10	R.Federer	Deuce	32	31	63	0.51	0.49	24	24	0.75	0.77	0.051	0.821
23-Jun-10	I.Bozoljac	Ad	31	27	58	0.53	0.47	25	20	0.81	0.74	0.358	0.549
23-Jun-10	I.Bozoljac	Deuce	38	23	61	0.62	0.38	28	17	0.74	0.74	0.000	0.984
25-Jun-10	R.Federer	Ad	17	11	28	0.61	0.39	15	8	0.88	0.73	1.095	0.295
25-Jun-10	R.Federer	Deuce	17	14	31	0.55	0.45	14	13	0.82	0.93	0.754	0.385
25-Jun-10	A.Clement	Ad	19	12	31	0.61	0.39	6	9	0.32	0.75	5.552	0.018*
25-Jun-10	A.Clement	Deuce	17	17	34	0.50	0.50	10	13	0.59	0.76	1.209	0.271
26-Jun-10	P.Petzschner	Ad	24	46	70	0.34	0.66	19	26	0.79	0.57	3.523	0.060**
26-Jun-10	P.Petzschner	Deuce	34	44	78	0.44	0.56	25	28	0.74	0.64	0.862	0.353
26-Jun-10	R.Nadal	Ad	34	17	51	0.67	0.33	32	15	0.94	0.88	0.543	0.461
26-Jun-10	R.Nadal	Deuce	42	12	54	0.78	0.22	28	8	0.67	0.67	0.000	0.999
28-Jun-10	N.Djokovic	Ad	23	27	50	0.46	0.54	16	20	0.70	0.74	0.125	0.723
28-Jun-10	N.Djokovic	Deuce	27	17	44	0.61	0.39	18	15	0.67	0.88	2.588	0.107
28-Jun-10	L.Hewitt	Ad	25	24	49	0.51	0.49	20	16	0.80	0.67	1.117	0.290
28-Jun-10	L.Hewitt	Deuce	13	32	45	0.29	0.71	10	22	0.77	0.69	0.301	0.583
28-Jun-10	R.Federer	Ad	9	17	26	0.35	0.65	7	12	0.78	0.71	0.155	0.694
28-Jun-10	R.Federer	Deuce	14	16	30	0.47	0.53	13	14	0.93	0.88	0.238	0.625
28-Jun-10	J.Melzer	Ad	24	11	35	0.69	0.31	14	6	0.58	0.55	0.044	0.833
28-Jun-10	J.Melzer	Deuce	27	10	37	0.73	0.27	11	7	0.41	0.70	2.501	0.113
30-Jun-10	R.Federer	Ad	32	22	54	0.59	0.41	19	13	0.59	0.59	0.000	0.983
30-Jun-10	R.Federer	Deuce	26	30	56	0.46	0.54	18	22	0.69	0.73	0.115	0.734
30-Jun-10	T.Berdych	Ad	32	20	52	0.62	0.38	18	16	0.56	0.80	3.067	0.079**
30-Jun-10	T.Berdych	Deuce	26	23	49	0.53	0.47	17	17	0.65	0.74	0.418	0.517
30-Jun-10	R.Soderling	Ad	18	22	40	0.45	0.55	12	15	0.67	0.68	0.010	0.918
30-Jun-10	R.Soderling	Deuce	23	23	46	0.50	0.50	20	16	0.87	0.70	2.044	0.152
30-Jun-10	R.Nadal	Ad	19	29	48	0.40	0.60	11	27	0.58	0.93	8.628	0.003*
30-Jun-10	R.Nadal	Deuce	32	18	50	0.64	0.36	20	12	0.63	0.67	0.087	0.768
30-Jun-10	J.Tsonga	Ad	30	34	64	0.47	0.53	23	21	0.77	0.62	1.647	0.199
30-Jun-10	J.Tsonga	Deuce	36	27	63	0.57	0.43	20	19	0.56	0.70	1.436	0.230
30-Jun-10	A.Murray	Ad	38	18	56	0.68	0.32	29	14	0.76	0.78	0.015	0.903
30-Jun-10	A.Murray	Deuce	35	20	55	0.64	0.36	26	15	0.74	0.75	0.003	0.953
02-Jul-10	A.Murray	Ad	11	29	40	0.28	0.73	8	19	0.73	0.66	0.189	0.663
02-Jul-10	A.Murray	Deuce	19	25	44	0.43	0.57	15	20	0.79	0.80	0.007	0.931
02-Jul-10	R.Nadal	Ad	12	27	39	0.31	0.69	6	22	0.50	0.81	4.066	0.043*
02-Jul-10	R.Nadal	Deuce	22	20	42	0.52	0.48	15	15	0.68	0.75	0.239	0.625
02-Jul-10	T.Berdych	Ad	23	18	41	0.56	0.44	15	15	0.65	0.83	1.688	0.193
02-Jul-10	T.Berdych	Deuce	20	22	42	0.48	0.52	15	17	0.75	0.77	0.030	0.862
02-Jul-10	N.Djokovic	Ad	18	13	31	0.58	0.42	13	12	0.72	0.92	1.951	0.162
02-Jul-10	N.Djokovic	Deuce	20	16	36	0.56	0.44	15	8	0.75	0.50	2.408	0.120
04-Jul-10	T.Berdych	Ad	8	26	34	0.24	0.76	7	15	0.88	0.58	2.380	0.122
04-Jul-10	T.Berdych	Deuce	6	34	40	0.15	0.85	4	22	0.67	0.65	0.009	0.926
04-Jul-10	R.Nadal	Ad	25	10	35	0.71	0.29	22	9	0.88	0.90	0.028	0.866
04-Jul-10	R.Nadal	Deuce	31	9	40	0.78	0.23	18	7	0.58	0.78	1.157	0.282
21-Jun-11	M.Kukushkin	Ad	26	13	39	0.67	0.33	12	9	0.46	0.69	1.857	0.172
21-Jun-11	M.Kukushkin	Deuce	23	22	45	0.51	0.49	14	17	0.61	0.77	1.412	0.234
21-Jun-11	R.Federer	Ad	24	13	37	0.65	0.35	21	9	0.88	0.69	1.835	0.175
21-Jun-11	R.Federer	Deuce	17	20	37	0.46	0.54	17	18	1.00	0.90	1.797	0.180
23-Jun-11	A.Mannarino	Ad	20	10	30	0.67	0.33	10	6	0.50	0.60	0.268	0.604
23-Jun-11	A.Mannarino	Deuce	24	10	34	0.71	0.29	15	4	0.63	0.40	1.449	0.228
23-Jun-11	R.Federer	Ad	15	16	31	0.48	0.52	13	14	0.87	0.88	0.005	0.944
23-Jun-11	R.Federer	Deuce	15	16	31	0.48	0.52	13	11	0.87	0.69	1.422	0.233
25-Jun-11	M.Baghdatis	Ad	27	22	49	0.55	0.45	17	15	0.63	0.68	0.146	0.702
25-Jun-11	M.Baghdatis	Deuce	21	31	52	0.40	0.60	12	20	0.57	0.65	0.288	0.591
25-Jun-11	N.Djokovic	Ad	19	24	43	0.44	0.56	13	19	0.68	0.79	0.643	0.422
25-Jun-11	N.Djokovic	Deuce	26	26	52	0.50	0.50	18	20	0.69	0.77	0.391	0.531
25-Jun-11	D.Nalbandian	Ad	21	17	38	0.55	0.45	12	10	0.57	0.59	0.011	0.916
25-Jun-11	D.Nalbandian	Deuce	19	13	32	0.59	0.41	13	8	0.68	0.62	0.162	0.687
25-Jun-11	R.Federer	Ad	24	13	37	0.65	0.35	15	9	0.63	0.69	0.168	0.682
25-Jun-11	R.Federer	Deuce	26	13	39	0.67	0.33	23	9	0.88	0.69	2.176	0.140
27-Jun-11	A.Murray	Ad	20	15	35	0.57	0.43	17	9	0.85	0.60	2.804	0.094**
27-Jun-11	A.Murray	Deuce	21	14	35	0.60	0.40	17	10	0.81	0.71	0.432	0.510
27-Jun-11	R.Gasquet	Ad	20	17	37	0.54	0.46	13	11	0.65	0.65	0.000	0.985
27-Jun-11	R.Gasquet	Deuce	19	22	41	0.46	0.54	12	16	0.63	0.73	0.431	0.511
27-Jun-11	M.Llodra	Ad	20	13	33	0.61	0.39	13	7	0.65	0.54	0.411	0.521
27-Jun-11	M.Llodra	Deuce	19	13	32	0.59	0.41	13	10	0.68	0.77	0.276	0.599
27-Jun-11	N.Djokovic	Ad	11	19	30	0.37	0.63	9	15	0.82	0.79	0.036	0.849
27-Jun-11	N.Djokovic	Deuce	12	19	31	0.39	0.61	11	14	0.92	0.74	1.524	0.217
27-Jun-11	M.Youzhny	Ad	42	19	61	0.69	0.31	22	11	0.52	0.58	0.160	0.689
27-Jun-11	M.Youzhny	Deuce	36	24	60	0.60	0.40	24	15	0.67	0.63	0.110	0.740
27-Jun-11	R.Federer	Ad	30	18	48	0.63	0.38	26	13	0.87	0.72	1.541	0.214
27-Jun-11	R.Federer	Deuce	19	25	44	0.43	0.57	12	20	0.63	0.80	1.544	0.214
29-Jun-11	R.Nadal	Ad	26	12	38	0.68	0.32	18	10	0.69	0.83	0.842	0.358
29-Jun-11	R.Nadal	Deuce	18	11	29	0.62	0.38	13	10	0.72	0.91	1.453	0.228
29-Jun-11	M.Fish	Ad	28	23	51	0.55	0.45	14	16	0.50	0.70	1.996	0.157

Continued on next page

Table A.3 – *Continued from previous page*

			Serves			Mixture		Won		Win rates		Individual test	
Year	Server	Court	L	R	Total	L	R	L	R	L	R	Pearson	p-value
29-Jun-11	M.Fish	Deuce	25	24	49	0.51	0.49	16	12	0.64	0.50	0.980	0.322
29-Jun-11	J.Tsonga	Ad	44	18	62	0.71	0.29	31	15	0.70	0.83	1.107	0.292
29-Jun-11	J.Tsonga	Deuce	40	24	64	0.63	0.38	28	16	0.70	0.67	0.078	0.780
29-Jun-11	R.Federer	Ad	34	22	56	0.61	0.39	25	16	0.74	0.73	0.004	0.947
29-Jun-11	R.Federer	Deuce	30	32	62	0.48	0.52	23	24	0.77	0.75	0.023	0.878
01-Jul-11	R.Nadal	Ad	13	22	35	0.37	0.63	10	19	0.77	0.86	0.513	0.473
01-Jul-11	R.Nadal	Deuce	28	15	43	0.65	0.35	25	9	0.89	0.60	5.062	0.024*
01-Jul-11	A.Murray	Ad	12	36	48	0.25	0.75	8	22	0.67	0.61	0.119	0.730
01-Jul-11	A.Murray	Deuce	22	27	49	0.45	0.55	21	13	0.95	0.48	12.772	0.000*
01-Jul-11	J.Tsonga	Ad	35	23	58	0.60	0.40	18	20	0.51	0.87	7.754	0.005*
01-Jul-11	J.Tsonga	Deuce	27	30	57	0.47	0.53	19	17	0.70	0.57	1.147	0.284
01-Jul-11	N.Djokovic	Ad	28	13	41	0.68	0.32	25	7	0.89	0.54	6.508	0.010*
01-Jul-11	N.Djokovic	Deuce	30	27	57	0.53	0.47	22	23	0.73	0.85	1.201	0.273
03-Jul-11	R.Nadal	Ad	12	19	31	0.39	0.61	11	12	0.92	0.63	3.122	0.077**
03-Jul-11	R.Nadal	Deuce	29	8	37	0.78	0.22	20	4	0.69	0.50	0.990	0.319
03-Jul-11	N.Djokovic	Ad	14	24	38	0.37	0.63	11	13	0.79	0.54	2.263	0.132
03-Jul-11	N.Djokovic	Deuce	16	27	43	0.37	0.63	11	22	0.69	0.81	0.912	0.339
25-Jun-12	R.Federer	Ad	13	11	24	0.54	0.46	10	4	0.77	0.36	4.033	0.044*
25-Jun-12	R.Federer	Deuce	16	12	28	0.57	0.43	14	10	0.88	0.83	0.097	0.755
25-Jun-12	A.Ramos	Ad	22	8	30	0.73	0.27	8	2	0.36	0.25	0.341	0.559
25-Jun-12	A.Ramos	Deuce	18	7	25	0.72	0.28	11	3	0.61	0.43	0.682	0.409
27-Jun-12	R.Federer	Ad	15	9	24	0.63	0.38	10	8	0.67	0.89	1.481	0.223
27-Jun-12	R.Federer	Deuce	19	9	28	0.68	0.32	17	8	0.89	0.89	0.002	0.962
27-Jun-12	F.Fognini	Ad	17	10	27	0.63	0.37	10	5	0.59	0.50	0.199	0.655
27-Jun-12	F.Fognini	Deuce	22	10	32	0.69	0.31	9	7	0.41	0.70	2.327	0.127
29-Jun-12	R.Federer	Ad	37	25	62	0.60	0.40	26	19	0.70	0.76	0.246	0.619
29-Jun-12	R.Federer	Deuce	33	39	72	0.46	0.54	26	31	0.79	0.79	0.005	0.941
29-Jun-12	J.Benneteau	Ad	39	26	65	0.60	0.40	28	18	0.72	0.69	0.050	0.823
29-Jun-12	J.Benneteau	Deuce	57	20	77	0.74	0.26	36	13	0.63	0.65	0.022	0.882
02-Jul-12	R.Federer	Ad	26	22	48	0.54	0.46	21	12	0.81	0.55	3.814	0.050**
02-Jul-12	R.Federer	Deuce	23	29	52	0.44	0.56	16	21	0.70	0.72	0.051	0.821
02-Jul-12	X.Malisse	Ad	30	20	50	0.60	0.40	21	9	0.70	0.45	3.125	0.077**
02-Jul-12	X.Malisse	Deuce	44	11	55	0.80	0.20	23	7	0.52	0.64	0.458	0.498
04-Jul-12	R.Federer	Ad	17	10	27	0.63	0.37	12	9	0.71	0.90	1.373	0.241
04-Jul-12	R.Federer	Deuce	10	19	29	0.34	0.66	6	15	0.60	0.79	1.177	0.277
04-Jul-12	M.Youzhny	Ad	15	16	31	0.48	0.52	8	9	0.53	0.56	0.027	0.870
04-Jul-12	M.Youzhny	Deuce	21	13	34	0.62	0.38	10	8	0.48	0.62	0.624	0.429
04-Jul-12	D.Ferrer	Ad	33	41	74	0.45	0.55	16	33	0.48	0.80	8.371	0.003*
04-Jul-12	D.Ferrer	Deuce	24	31	55	0.44	0.56	14	24	0.58	0.77	2.307	0.128
04-Jul-12	A.Murray	Ad	38	27	65	0.58	0.42	34	20	0.89	0.74	2.663	0.102
04-Jul-12	A.Murray	Deuce	36	29	65	0.55	0.45	24	19	0.67	0.66	0.009	0.922
06-Jul-12	A.Murray	Ad	21	24	45	0.47	0.53	16	16	0.76	0.67	0.495	0.481
06-Jul-12	A.Murray	Deuce	30	24	54	0.56	0.44	22	18	0.73	0.75	0.019	0.889
06-Jul-12	J.Tsonga	Ad	42	17	59	0.71	0.29	27	12	0.64	0.71	0.215	0.643
06-Jul-12	J.Tsonga	Deuce	35	22	57	0.61	0.39	18	17	0.51	0.77	3.807	0.051**
06-Jul-12	N.Djokovic	Ad	25	20	45	0.56	0.44	13	15	0.52	0.75	2.501	0.113
06-Jul-12	N.Djokovic	Deuce	30	22	52	0.58	0.42	20	18	0.67	0.82	1.481	0.223
06-Jul-12	R.Federer	Ad	21	20	41	0.51	0.49	17	14	0.81	0.70	0.666	0.414
06-Jul-12	R.Federer	Deuce	22	23	45	0.49	0.51	17	19	0.77	0.83	0.200	0.654
08-Jul-12	R.Federer	Ad	35	25	60	0.58	0.42	22	17	0.63	0.68	0.170	0.680
08-Jul-12	R.Federer	Deuce	22	33	55	0.40	0.60	17	27	0.77	0.82	0.170	0.679
08-Jul-12	A.Murray	Ad	52	21	73	0.71	0.29	36	9	0.69	0.43	4.401	0.035*
08-Jul-12	A.Murray	Deuce	53	20	73	0.73	0.27	31	14	0.58	0.70	0.814	0.367
24-Jun-13	V.Hanescu	Ad	12	9	21	0.57	0.43	5	4	0.42	0.44	0.016	0.898
24-Jun-13	V.Hanescu	Deuce	13	11	24	0.54	0.46	8	6	0.62	0.55	0.120	0.729
24-Jun-13	R.Federer	Ad	9	11	20	0.45	0.55	8	10	0.89	0.91	0.022	0.880
24-Jun-13	R.Federer	Deuce	14	11	25	0.56	0.44	13	9	0.93	0.82	0.711	0.399
24-Jun-13	R.Nadal	Ad	29	17	46	0.63	0.37	18	13	0.62	0.76	1.012	0.314
24-Jun-13	R.Nadal	Deuce	40	10	50	0.80	0.20	24	5	0.60	0.50	0.328	0.566
24-Jun-13	S.Darcis	Ad	18	31	49	0.37	0.63	10	20	0.56	0.65	0.385	0.534
24-Jun-13	S.Darcis	Deuce	20	34	54	0.37	0.63	17	22	0.85	0.65	2.585	0.107
25-Jun-13	N.Djokovic	Ad	22	17	39	0.56	0.44	21	14	0.95	0.82	1.788	0.181
25-Jun-13	N.Djokovic	Deuce	19	21	40	0.48	0.53	13	11	0.68	0.52	1.069	0.301
25-Jun-13	F.Mayer	Ad	24	17	41	0.59	0.41	14	10	0.58	0.59	0.001	0.974
25-Jun-13	F.Mayer	Deuce	28	17	45	0.62	0.38	17	14	0.61	0.82	2.311	0.128
25-Jun-13	D.Ferrer	Ad	20	18	38	0.53	0.47	18	14	0.90	0.78	1.064	0.302
25-Jun-13	D.Ferrer	Deuce	11	16	27	0.41	0.59	9	13	0.82	0.81	0.001	0.970
25-Jun-13	M.Alund	Ad	31	15	46	0.67	0.33	22	11	0.71	0.73	0.028	0.867
25-Jun-13	M.Alund	Deuce	22	14	36	0.61	0.39	10	9	0.45	0.64	1.217	0.269
25-Jun-13	A.Ramos	Ad	19	6	25	0.76	0.24	13	4	0.68	0.67	0.006	0.935
25-Jun-13	A.Ramos	Deuce	25	14	39	0.64	0.36	12	8	0.48	0.57	0.300	0.583
25-Jun-13	J.Del Potro	Ad	8	25	33	0.24	0.76	6	21	0.75	0.84	0.330	0.565
25-Jun-13	J.Del Potro	Deuce	18	12	30	0.60	0.40	14	9	0.78	0.75	0.031	0.860
26-Jun-13	J-W.Tsonga	Ad	22	13	35	0.63	0.37	13	8	0.59	0.62	0.020	0.886
26-Jun-13	J-W.Tsonga	Deuce	17	17	34	0.50	0.50	16	10	0.94	0.59	5.885	0.015*
26-Jun-13	E.Gulbis	Ad	15	14	29	0.52	0.48	10	11	0.67	0.79	0.514	0.473
26-Jun-13	E.Gulbis	Deuce	13	18	31	0.42	0.58	11	12	0.85	0.67	1.270	0.259
26-Jun-13	F.Verdasco	Ad	28	13	41	0.68	0.32	24	12	0.86	0.92	0.360	0.548
26-Jun-13	F.Verdasco	Deuce	22	20	42	0.52	0.48	19	15	0.86	0.75	0.877	0.348

Continued on next page

Table A.3 – *Continued from previous page*

			Serves			Mixture		Won		Win rates		Individual test	
Year	Server	Court	L	R	Total	L	R	L	R	L	R	Pearson	p-value
26-Jun-13	J.Benneteau	Ad	21	26	47	0.45	0.55	14	21	0.67	0.81	1.215	0.270
26-Jun-13	J.Benneteau	Deuce	27	25	52	0.52	0.48	19	14	0.70	0.56	1.156	0.282
27-Jun-13	N.Djokovic	Ad	18	14	32	0.56	0.44	14	11	0.78	0.79	0.003	0.957
27-Jun-13	N.Djokovic	Deuce	17	20	37	0.46	0.54	15	19	0.88	0.95	0.564	0.452
27-Jun-13	B.Reynolds	Ad	16	22	38	0.42	0.58	10	12	0.63	0.55	0.240	0.623
27-Jun-13	B.Reynolds	Deuce	15	12	27	0.56	0.44	13	9	0.87	0.75	0.601	0.438
27-Jun-13	R.Gasquet	Ad	16	23	39	0.41	0.59	11	18	0.69	0.78	0.448	0.503
27-Jun-13	R.Gasquet	Deuce	28	20	48	0.58	0.42	21	13	0.75	0.65	0.565	0.452
27-Jun-13	G.Soeda	Ad	26	20	46	0.57	0.43	17	8	0.65	0.40	2.936	0.086**
27-Jun-13	G.Soeda	Deuce	26	22	48	0.54	0.46	16	13	0.62	0.59	0.030	0.862
27-Jun-13	J.Levine	Ad	25	16	41	0.61	0.39	14	13	0.56	0.81	2.766	0.096**
27-Jun-13	J.Levine	Deuce	24	16	40	0.60	0.40	12	11	0.50	0.69	1.381	0.239
27-Jun-13	J.Del Potro	Ad	7	29	36	0.19	0.81	3	24	0.43	0.83	4.788	0.028*
27-Jun-13	J.Del Potro	Deuce	20	9	29	0.69	0.31	15	7	0.75	0.78	0.026	0.871
28-Jun-13	J.Melzer	Ad	31	11	42	0.74	0.26	19	10	0.61	0.91	3.333	0.067**
28-Jun-13	J.Melzer	Deuce	24	20	44	0.55	0.45	17	11	0.71	0.55	1.182	0.276
28-Jun-13	S.Stakhovsky	Ad	17	13	30	0.57	0.43	11	8	0.65	0.62	0.032	0.858
28-Jun-13	S.Stakhovsky	Deuce	17	19	36	0.47	0.53	13	13	0.76	0.68	0.290	0.590
01-Jul-13	J.Janowicz	Ad	27	29	56	0.48	0.52	20	23	0.74	0.79	0.215	0.642
01-Jul-13	J.Janowicz	Deuce	30	29	59	0.51	0.49	23	24	0.77	0.83	0.338	0.561
01-Jul-13	J.Melzer	Ad	44	14	58	0.76	0.24	35	11	0.80	0.79	0.006	0.937
01-Jul-13	J.Melzer	Deuce	27	24	51	0.53	0.47	15	18	0.56	0.75	2.104	0.146
01-Jul-13	M.Youzhny	Ad	27	13	40	0.68	0.33	14	7	0.52	0.54	0.014	0.905
01-Jul-13	M.Youzhny	Deuce	22	18	40	0.55	0.45	11	12	0.50	0.67	1.125	0.288
01-Jul-13	A.Murray	Ad	28	14	42	0.67	0.33	16	11	0.57	0.79	1.867	0.171
01-Jul-13	A.Murray	Deuce	35	13	48	0.73	0.27	32	7	0.91	0.54	8.789	0.003*
03-Jul-13	N.Djokovic	Ad	20	20	40	0.50	0.50	15	13	0.75	0.65	0.476	0.490
03-Jul-13	N.Djokovic	Deuce	25	13	38	0.66	0.34	21	8	0.84	0.62	2.387	0.122
03-Jul-13	T.Berdych	Ad	22	12	34	0.65	0.35	16	5	0.73	0.42	3.172	0.074**
03-Jul-13	T.Berdych	Deuce	15	22	37	0.41	0.59	9	15	0.60	0.68	0.262	0.608
03-Jul-13	F.Verdasco	Ad	45	21	66	0.68	0.32	30	14	0.67	0.67	0.000	0.999
03-Jul-13	F.Verdasco	Deuce	29	34	63	0.46	0.54	20	21	0.69	0.62	0.357	0.550
03-Jul-13	A.Murray	Ad	9	39	48	0.19	0.81	7	27	0.78	0.69	0.259	0.611
03-Jul-13	A.Murray	Deuce	32	30	62	0.52	0.48	28	17	0.88	0.57	7.397	0.006*
05-Jul-13	N.Djokovic	Ad	47	27	74	0.64	0.36	37	20	0.79	0.74	0.209	0.647
05-Jul-13	N.Djokovic	Deuce	60	24	84	0.71	0.29	37	20	0.62	0.83	3.690	0.054**
05-Jul-13	J.Del Potro	Ad	32	46	78	0.41	0.59	19	32	0.59	0.70	0.866	0.352
05-Jul-13	J.Del Potro	Deuce	45	15	60	0.75	0.25	32	12	0.71	0.80	0.455	0.500
05-Jul-13	J.Janowicz	Ad	33	22	55	0.60	0.40	23	14	0.70	0.64	0.220	0.638
05-Jul-13	J.Janowicz	Deuce	32	24	56	0.57	0.43	24	16	0.75	0.67	0.467	0.494
05-Jul-13	A.Murray	Ad	25	22	47	0.53	0.47	19	15	0.76	0.68	0.357	0.549
05-Jul-13	A.Murray	Deuce	30	20	50	0.60	0.40	22	17	0.73	0.85	0.952	0.329
07-Jul-13	N.Djokovic	Ad	28	18	46	0.61	0.39	18	8	0.64	0.44	1.755	0.185
07-Jul-13	N.Djokovic	Deuce	24	19	43	0.56	0.44	13	11	0.54	0.58	0.060	0.806
07-Jul-13	A.Murray	Ad	27	18	45	0.60	0.40	16	13	0.59	0.72	0.792	0.373
07-Jul-13	A.Murray	Deuce	27	19	46	0.59	0.41	17	12	0.63	0.63	0.000	0.989
23-Jun-14	G.Dimitrov	Ad	17	17	34	0.50	0.50	13	12	0.76	0.71	0.151	0.697
23-Jun-14	G.Dimitrov	Deuce	19	17	36	0.53	0.47	17	13	0.89	0.76	1.092	0.295
23-Jun-14	R.Harrison	Ad	19	17	36	0.53	0.47	13	9	0.68	0.53	0.905	0.341
23-Jun-14	R.Harrison	Deuce	20	17	37	0.54	0.46	15	6	0.75	0.35	5.903	0.015*
23-Jun-14	A.Murray	Ad	15	21	36	0.42	0.58	10	17	0.67	0.81	0.952	0.329
23-Jun-14	A.Murray	Deuce	27	12	39	0.69	0.31	22	11	0.81	0.92	0.662	0.415
23-Jun-14	D.Goffin	Ad	22	14	36	0.61	0.39	13	8	0.59	0.57	0.013	0.907
23-Jun-14	D.Goffin	Deuce	25	18	43	0.58	0.42	17	12	0.68	0.67	0.008	0.926
24-Jun-14	P.Lorenzi	Ad	40	6	46	0.87	0.13	22	2	0.55	0.33	0.982	0.321
24-Jun-14	P.Lorenzi	Deuce	32	7	39	0.82	0.18	20	0	0.63	0.00	8.980	0.002*
24-Jun-14	R.Federer	Ad	14	15	29	0.48	0.52	10	15	0.71	1.00	4.971	0.025*
24-Jun-14	R.Federer	Deuce	16	14	30	0.53	0.47	13	8	0.81	0.57	2.066	0.150
24-Jun-14	S.Wawrinka	Ad	17	16	33	0.52	0.48	10	14	0.59	0.88	3.417	0.064**
24-Jun-14	S.Wawrinka	Deuce	25	10	35	0.71	0.29	20	10	0.80	1.00	2.333	0.126
24-Jun-14	J.Sousa	Ad	26	6	32	0.81	0.19	16	6	0.62	1.00	3.357	0.066**
24-Jun-14	J.Sousa	Deuce	12	21	33	0.36	0.64	7	13	0.58	0.62	0.041	0.839
25-Jun-14	G.Dimitrov	Ad	20	9	29	0.69	0.31	17	8	0.85	0.89	0.079	0.778
25-Jun-14	G.Dimitrov	Deuce	16	17	33	0.48	0.52	15	14	0.94	0.82	1.005	0.316
25-Jun-14	L.Saville	Ad	26	7	33	0.79	0.21	17	4	0.65	0.57	0.162	0.687
25-Jun-14	L.Saville	Deuce	24	11	35	0.69	0.31	15	7	0.63	0.64	0.004	0.948
25-Jun-14	A.Murray	Ad	8	18	26	0.31	0.69	8	10	1.00	0.56	5.136	0.023*
25-Jun-14	A.Murray	Deuce	10	15	25	0.40	0.60	8	12	0.80	0.80	0.000	0.999
25-Jun-14	B.Rola	Ad	14	3	17	0.82	0.18	7	1	0.50	0.33	0.275	0.599
25-Jun-14	B.Rola	Deuce	15	8	23	0.65	0.35	4	3	0.27	0.38	0.289	0.590
26-Jun-14	G.Muller	Ad	20	14	34	0.59	0.41	14	10	0.70	0.71	0.008	0.928
26-Jun-14	G.Muller	Deuce	30	12	42	0.71	0.29	20	8	0.67	0.67	0.000	0.999
26-Jun-14	R.Federer	Ad	19	13	32	0.59	0.41	16	12	0.84	0.92	0.463	0.496
26-Jun-14	R.Federer	Deuce	16	16	32	0.50	0.50	13	15	0.81	0.94	1.143	0.285
26-Jun-14	L.Rosol	Ad	19	29	48	0.40	0.60	15	18	0.79	0.62	1.522	0.217
26-Jun-14	L.Rosol	Deuce	17	34	51	0.33	0.67	15	20	0.88	0.59	4.554	0.032*
26-Jun-14	R.Nadal	Ad	31	14	45	0.69	0.31	24	7	0.77	0.50	3.383	0.065**
26-Jun-14	R.Nadal	Deuce	36	13	49	0.73	0.27	28	11	0.78	0.85	0.275	0.600
26-Jun-14	S.Wawrinka	Ad	35	24	59	0.59	0.41	24	16	0.69	0.67	0.024	0.877

Continued on next page

Table A.3 – *Continued from previous page*

			Serves			Mixture		Won		Win rates		Individual test	
Year	Server	Court	L	R	Total	L	R	L	R	L	R	Pearson	p-value
26-Jun-14	S.Wawrinka	Deuce	39	29	68	0.57	0.43	28	19	0.72	0.66	0.307	0.579
26-Jun-14	Y-H.Lu	Ad	32	24	56	0.57	0.43	22	13	0.69	0.54	1.244	0.264
26-Jun-14	Y-H.Lu	Deuce	20	33	53	0.38	0.62	10	22	0.50	0.67	1.446	0.229
27-Jun-14	N.Djokovic	Ad	20	13	33	0.61	0.39	10	10	0.50	0.77	2.392	0.121
27-Jun-14	N.Djokovic	Deuce	17	17	34	0.50	0.50	14	11	0.82	0.65	1.360	0.243
27-Jun-14	G.Simon	Ad	23	23	46	0.50	0.50	10	12	0.43	0.52	0.348	0.554
27-Jun-14	G.Simon	Deuce	20	22	42	0.48	0.52	12	15	0.60	0.68	0.305	0.580
28-Jun-14	S.Giraldo	Ad	24	13	37	0.65	0.35	13	4	0.54	0.31	1.859	0.172
28-Jun-14	S.Giraldo	Deuce	21	10	31	0.68	0.32	16	7	0.76	0.70	0.136	0.712
28-Jun-14	R.Federer	Ad	19	14	33	0.58	0.42	18	11	0.95	0.79	1.977	0.159
28-Jun-14	R.Federer	Deuce	19	15	34	0.56	0.44	13	9	0.68	0.60	0.260	0.609
28-Jun-14	M.Kukushkin	Ad	23	33	56	0.41	0.59	16	15	0.70	0.45	3.188	0.074**
28-Jun-14	M.Kukushkin	Deuce	22	32	54	0.41	0.59	15	17	0.68	0.53	1.224	0.268
28-Jun-14	R.Nadal	Ad	12	16	28	0.43	0.57	8	13	0.67	0.81	0.778	0.377
28-Jun-14	R.Nadal	Deuce	29	9	38	0.76	0.24	21	8	0.72	0.89	1.031	0.309
30-Jun-14	S.Wawrinka	Ad	18	11	29	0.62	0.38	13	10	0.72	0.91	1.453	0.228
30-Jun-14	S.Wawrinka	Deuce	20	9	29	0.69	0.31	15	7	0.75	0.78	0.026	0.871
30-Jun-14	D.Istomin	Ad	13	14	27	0.48	0.52	7	9	0.54	0.64	0.304	0.581
30-Jun-14	D.Istomin	Deuce	19	13	32	0.59	0.41	13	7	0.68	0.54	0.700	0.402
02-Jul-14	A.Murray	Ad	18	16	34	0.53	0.47	11	11	0.61	0.69	0.216	0.641
02-Jul-14	A.Murray	Deuce	20	15	35	0.57	0.43	13	8	0.65	0.53	0.486	0.485
02-Jul-14	G.Dimitrov	Ad	17	21	38	0.45	0.55	14	13	0.82	0.62	1.910	0.166
02-Jul-14	G.Dimitrov	Deuce	17	20	37	0.46	0.54	12	15	0.71	0.75	0.091	0.763
02-Jul-14	S.Wawrinka	Ad	34	21	55	0.62	0.38	20	17	0.59	0.81	2.887	0.089**
02-Jul-14	S.Wawrinka	Deuce	34	26	60	0.57	0.43	21	22	0.62	0.85	3.789	0.051**
02-Jul-14	R.Federer	Ad	35	20	55	0.64	0.36	24	17	0.69	0.85	1.810	0.178
02-Jul-14	R.Federer	Deuce	33	19	52	0.63	0.37	26	14	0.79	0.74	0.177	0.674
04-Jul-14	N.Djokovic	Ad	35	27	62	0.56	0.44	22	18	0.63	0.67	0.097	0.755
04-Jul-14	N.Djokovic	Deuce	25	37	62	0.40	0.60	18	29	0.72	0.78	0.331	0.565
04-Jul-14	G.Dimitrov	Ad	23	29	52	0.44	0.56	15	21	0.65	0.72	0.312	0.576
04-Jul-14	G.Dimitrov	Deuce	29	19	48	0.60	0.40	20	17	0.69	0.89	2.733	0.098**
04-Jul-14	R.Federer	Ad	19	14	33	0.58	0.42	16	11	0.84	0.79	0.172	0.678
04-Jul-14	R.Federer	Deuce	18	18	36	0.50	0.50	12	16	0.67	0.89	2.571	0.108
04-Jul-14	M.Raonic	Ad	27	10	37	0.73	0.27	19	6	0.70	0.60	0.358	0.549
04-Jul-14	M.Raonic	Deuce	15	19	34	0.44	0.56	10	16	0.67	0.84	1.434	0.231
06-Jul-14	N.Djokovic	Ad	35	29	64	0.55	0.45	22	20	0.63	0.69	0.262	0.608
06-Jul-14	N.Djokovic	Deuce	47	28	75	0.63	0.37	36	22	0.77	0.79	0.039	0.843
06-Jul-14	R.Federer	Ad	40	36	76	0.53	0.47	30	25	0.75	0.69	0.292	0.588
06-Jul-14	R.Federer	Deuce	35	52	87	0.40	0.60	22	38	0.63	0.73	1.021	0.312
Sum			8,145	6,513	14,658	-	-	5,689	4,579	-	-	407.3	-
Mean			25.5	20.4	45.8	0.56	0.44	17.8	14.3	0.703	0.701	-	-

Table A.4
Prediction 2: Full sample results.

			Serves			Runs test			
Year	Server	Court	L	R	Total	Runs (r)	$F(r-1)$	$F(r)$	$U[F(r-1), F(r)]$
24-Jun-09	G.Garcia-Lopez	Ad	15	11	26	16	0.77	0.875	0.066
24-Jun-09	G.Garcia-Lopez	Deuce	20	11	31	19	0.907	0.957	0.97
24-Jun-09	R.Federer	Ad	20	13	33	15	0.201	0.32	0.317
24-Jun-09	R.Federer	Deuce	21	13	34	14	0.094	0.172	0.463
26-Jun-09	P.Kohlschreiber	Ad	27	22	49	27	0.642	0.744	0.265
26-Jun-09	P.Kohlschreiber	Deuce	23	25	48	22	0.156	0.236	0.438
26-Jun-09	R.Federer	Ad	30	25	55	28	0.416	0.524	0.63
26-Jun-09	R.Federer	Deuce	31	27	58	33	0.758	0.833	0.89
27-Jun-09	S.Wawrinka	Ad	24	29	53	30	0.734	0.817	0.988
27-Jun-09	S.Wawrinka	Deuce	16	38	54	24	0.497	0.627	0.188
27-Jun-09	J.Levine	Ad	37	27	64	34	0.629	0.722	0.826
27-Jun-09	J.Levine	Deuce	24	45	69	33	0.52	0.625	0.84
29-Jun-09	A.Murray	Ad	39	22	61	25	0.097	0.154	0.15
29-Jun-09	A.Murray	Deuce	29	37	66	34	0.498	0.597	0.532
29-Jun-09	S.Wawrinka	Ad	38	34	72	42	0.863	0.909	0.875
29-Jun-09	S.Wawrinka	Deuce	38	38	76	29	0.007	0.014*	0.01
29-Jun-09	R.Soderling	Ad	22	17	39	22	0.668	0.778	0.755
29-Jun-09	R.Soderling	Deuce	14	20	34	21	0.862	0.926	0.906
29-Jun-09	R.Federer	Ad	24	20	44	18	0.05	0.091	0.073
29-Jun-09	R.Federer	Deuce	22	24	46	18	0.026	0.051	0.033
01-Jul-09	T.Haas	Ad	20	25	45	22	0.299	0.412	0.315
01-Jul-09	T.Haas	Deuce	33	21	54	29	0.702	0.793	0.755
01-Jul-09	N.Djokovic	Ad	28	18	46	26	0.791	0.869	0.835
01-Jul-09	N.Djokovic	Deuce	25	24	49	16	0.001	0.004*	0.002
01-Jul-09	L.Hewitt	Ad	37	42	79	48	0.948	0.968	0.951
01-Jul-09	L.Hewitt	Deuce	25	48	73	32	0.266	0.359	0.268
01-Jul-09	A.Roddick	Ad	44	33	77	37	0.301	0.388	0.349
01-Jul-09	A.Roddick	Deuce	49	25	74	31	0.172	0.247	0.223
01-Jul-09	I.Karlovic	Ad	19	15	34	20	0.73	0.833	0.77
01-Jul-09	I.Karlovic	Deuce	22	16	38	19	0.364	0.496	0.487
01-Jul-09	R.Federer	Ad	20	15	35	13	0.023	0.051	0.028
01-Jul-09	R.Federer	Deuce	28	12	40	18	0.454	0.605	0.483
03-Jul-09	T.Haas	Ad	27	17	44	24	0.7	0.802	0.82
03-Jul-09	T.Haas	Deuce	37	12	49	17	0.151	0.261	0.92
03-Jul-09	R.Federer	Ad	25	14	39	17	0.193	0.304	0.318
03-Jul-09	R.Federer	Deuce	20	18	38	22	0.695	0.8	0.171
03-Jul-09	A.Roddick	Ad	29	32	61	34	0.704	0.786	0.717
03-Jul-09	A.Roddick	Deuce	37	25	62	29	0.266	0.36	0.225
03-Jul-09	A.Murray	Ad	45	18	63	20	0.012	0.026**	0.278
03-Jul-09	A.Murray	Deuce	45	17	62	25	0.351	0.477	0.783
05-Jul-09	A.Roddick	Ad	41	53	94	39	0.032	0.051	0.718
05-Jul-09	A.Roddick	Deuce	73	24	97	37	0.431	0.541	0.278
05-Jul-09	R.Federer	Ad	47	33	80	40	0.474	0.566	0.016
05-Jul-09	R.Federer	Deuce	45	42	87	43	0.336	0.418	0.462
21-Jun-10	R.Federer	Ad	28	39	67	27	0.036	0.061	0.04
21-Jun-10	R.Federer	Deuce	30	39	69	26	0.01	0.018*	0.477
21-Jun-10	A.Falla	Ad	71	12	83	21	0.32	0.494	0.522
21-Jun-10	A.Falla	Deuce	72	10	82	21	0.846	0.939	0.377
23-Jun-10	R.Federer	Ad	31	26	57	27	0.226	0.315	0.654
23-Jun-10	R.Federer	Deuce	32	31	63	30	0.223	0.306	0.216
23-Jun-10	I.Bozoljac	Ad	31	27	58	29	0.358	0.461	0.431
23-Jun-10	I.Bozoljac	Deuce	38	23	61	31	0.591	0.694	0.774
25-Jun-10	R.Federer	Ad	17	11	28	13	0.226	0.364	0.75
25-Jun-10	R.Federer	Deuce	17	14	31	14	0.146	0.246	0.574
25-Jun-10	A.Clement	Ad	19	12	31	18	0.755	0.859	0.707
25-Jun-10	A.Clement	Deuce	17	17	34	16	0.191	0.3	0.611
26-Jun-10	P.Petzschner	Ad	24	46	70	33	0.495	0.601	0.049
26-Jun-10	P.Petzschner	Deuce	34	44	78	38	0.333	0.421	0.014
26-Jun-10	R.Nadal	Ad	34	17	51	28	0.889	0.938	0.413
26-Jun-10	R.Nadal	Deuce	42	12	54	23	0.871	0.937	0.857
28-Jun-10	N.Djokovic	Ad	23	27	50	19	0.017	0.034**	0.033
28-Jun-10	N.Djokovic	Deuce	27	17	44	18	0.079	0.139	0.127
28-Jun-10	L.Hewitt	Ad	25	24	49	30	0.876	0.926	0.89
28-Jun-10	L.Hewitt	Deuce	13	32	45	21	0.645	0.77	0.714
28-Jun-10	R.Federer	Ad	9	17	26	11	0.156	0.286	0.195
28-Jun-10	R.Federer	Deuce	14	16	30	9	0.002	0.008*	0.002
28-Jun-10	J.Melzer	Ad	24	11	35	19	0.832	0.913	0.838
28-Jun-10	J.Melzer	Deuce	27	10	37	15	0.32	0.483	0.352
30-Jun-10	R.Federer	Ad	32	22	54	21	0.03	0.056	0.031
30-Jun-10	R.Federer	Deuce	26	30	56	23	0.042	0.073	0.055
30-Jun-10	T.Berdych	Ad	32	20	52	28	0.711	0.803	0.793
30-Jun-10	T.Berdych	Deuce	26	23	49	31	0.93	0.961	0.942
30-Jun-10	R.Soderling	Ad	18	22	40	27	0.967**	0.984	0.98
30-Jun-10	R.Soderling	Deuce	23	23	46	25	0.559	0.672	0.637
30-Jun-10	R.Nadal	Ad	19	29	48	28	0.86	0.917	0.864
30-Jun-10	R.Nadal	Deuce	32	18	50	26	0.674	0.777	0.675
30-Jun-10	J.Tsonga	Ad	30	34	64	27	0.053	0.086	0.062
30-Jun-10	J.Tsonga	Deuce	36	27	63	28	0.129	0.191	0.171

Continued on next page

Table A.4 – *Continued from previous page*

			Serves			Runs test			
Year	Server	Court	L	R	Total	Runs (r)	$F(r-1)$	$F(r)$	$U[F(r-1), F(r)]$
30-Jun-10	A.Murray	Ad	38	18	56	22	0.111	0.181	0.175
30-Jun-10	A.Murray	Deuce	35	20	55	33	0.962**	0.981	0.968
02-Jul-10	A.Murray	Ad	11	29	40	16	0.278	0.427	0.247
02-Jul-10	A.Murray	Deuce	19	25	44	20	0.168	0.257	0.29
02-Jul-10	R.Nadal	Ad	12	27	39	19	0.632	0.764	0.42
02-Jul-10	R.Nadal	Deuce	22	20	42	24	0.686	0.787	0.668
02-Jul-10	T.Berdych	Ad	23	18	41	19	0.193	0.293	0.377
02-Jul-10	T.Berdych	Deuce	20	22	42	23	0.568	0.686	0.182
02-Jul-10	N.Djokovic	Ad	18	13	31	13	0.088	0.164	0.712
02-Jul-10	N.Djokovic	Deuce	20	16	36	23	0.898	0.947	0.735
04-Jul-10	T.Berdych	Ad	8	26	34	13	0.359	0.551	0.198
04-Jul-10	T.Berdych	Deuce	6	34	40	11	0.325	0.576	0.648
04-Jul-10	R.Nadal	Ad	25	10	35	18	0.825	0.913	0.142
04-Jul-10	R.Nadal	Deuce	31	9	40	19	0.950**	0.982	0.934
21-Jun-11	M.Kukushkin	Ad	26	13	39	20	0.665	0.786	0.467
21-Jun-11	M.Kukushkin	Deuce	23	22	45	28	0.886	0.934	0.409
21-Jun-11	R.Federer	Ad	24	13	37	17	0.308	0.446	0.912
21-Jun-11	R.Federer	Deuce	17	20	37	17	0.166	0.264	0.977
23-Jun-11	A.Mannarino	Ad	20	10	30	16	0.687	0.818	0.343
23-Jun-11	A.Mannarino	Deuce	24	10	34	19	0.923	0.967	0.212
23-Jun-11	R.Federer	Ad	15	16	31	16	0.359	0.502	0.836
23-Jun-11	R.Federer	Deuce	15	16	31	13	0.072	0.137	0.279
25-Jun-11	M.Baghdatis	Ad	27	22	49	28	0.744	0.828	0.538
25-Jun-11	M.Baghdatis	Deuce	21	31	52	18	0.006	0.014*	0.403
25-Jun-11	N.Djokovic	Ad	19	24	43	21	0.296	0.412	0.92
25-Jun-11	N.Djokovic	Deuce	26	26	52	20	0.017	0.034**	0.878
25-Jun-11	D.Nalbandian	Ad	21	17	38	22	0.715	0.816	0.765
25-Jun-11	D.Nalbandian	Deuce	19	13	32	16	0.363	0.509	0.914
25-Jun-11	R.Federer	Ad	24	13	37	16	0.192	0.308	0.391
25-Jun-11	R.Federer	Deuce	26	13	39	13	0.016	0.038**	0.211
27-Jun-11	A.Murray	Ad	20	15	35	15	0.1	0.177	0.157
27-Jun-11	A.Murray	Deuce	21	14	35	22	0.907	0.953	0.924
27-Jun-11	R.Gasquet	Ad	20	17	37	18	0.264	0.384	0.276
27-Jun-11	R.Gasquet	Deuce	19	22	41	25	0.838	0.904	0.877
27-Jun-11	M.Llodra	Ad	20	13	33	19	0.74	0.845	0.842
27-Jun-11	M.Llodra	Deuce	19	13	32	18	0.654	0.779	0.666
27-Jun-11	N.Djokovic	Ad	11	19	30	20	0.966**	0.987	0.972
27-Jun-11	N.Djokovic	Deuce	12	19	31	15	0.32	0.467	0.415
27-Jun-11	M.Youzhny	Ad	42	19	61	32	0.904	0.946	0.923
27-Jun-11	M.Youzhny	Deuce	36	24	60	37	0.965**	0.981	0.974
27-Jun-11	R.Federer	Ad	30	18	48	25	0.622	0.733	0.698
27-Jun-11	R.Federer	Deuce	19	25	44	23	0.488	0.611	0.576
29-Jun-11	R.Nadal	Ad	26	12	38	14	0.066	0.132	0.129
29-Jun-11	R.Nadal	Deuce	18	11	29	15	0.475	0.633	0.597
29-Jun-11	M.Fish	Ad	28	23	51	26	0.414	0.527	0.472
29-Jun-11	M.Fish	Deuce	25	24	49	25	0.387	0.501	0.495
29-Jun-11	J.Tsonga	Ad	44	18	62	25	0.261	0.371	0.272
29-Jun-11	J.Tsonga	Deuce	40	24	64	23	0.011	0.021*	0.019
29-Jun-11	R.Federer	Ad	34	22	56	23	0.07	0.116	0.092
29-Jun-11	R.Federer	Deuce	30	32	62	26	0.048	0.08	0.057
01-Jul-11	R.Nadal	Ad	13	22	35	19	0.664	0.786	0.769
01-Jul-11	R.Nadal	Deuce	28	15	43	22	0.628	0.748	0.938
01-Jul-11	A.Murray	Ad	12	36	48	16	0.085	0.163	0.471
01-Jul-11	A.Murray	Deuce	22	27	49	27	0.642	0.744	0.077
01-Jul-11	J.Tsonga	Ad	35	23	58	30	0.581	0.685	0.725
01-Jul-11	J.Tsonga	Deuce	27	30	57	29	0.402	0.508	0.635
01-Jul-11	N.Djokovic	Ad	28	13	41	15	0.059	0.116	0.116
01-Jul-11	N.Djokovic	Deuce	30	27	57	29	0.402	0.508	0.711
03-Jul-11	R.Nadal	Ad	12	19	31	16	0.467	0.619	0.643
03-Jul-11	R.Nadal	Deuce	29	8	37	16	0.835	0.93	0.454
03-Jul-11	N.Djokovic	Ad	14	24	38	17	0.219	0.337	0.076
03-Jul-11	N.Djokovic	Deuce	16	27	43	18	0.117	0.195	0.475
25-Jun-12	R.Federer	Ad	13	11	24	14	0.596	0.747	0.473
25-Jun-12	R.Federer	Deuce	16	12	28	12	0.102	0.191	0.888
25-Jun-12	A.Ramos	Ad	22	8	30	13	0.455	0.643	0.321
25-Jun-12	A.Ramos	Deuce	18	7	25	13	0.766	0.892	0.125
27-Jun-12	R.Federer	Ad	15	9	24	10	0.109	0.217	0.776
27-Jun-12	R.Federer	Deuce	19	9	28	13	0.375	0.55	0.009
27-Jun-12	F.Fognini	Ad	17	10	27	14	0.484	0.649	0.408
27-Jun-12	F.Fognini	Deuce	22	10	32	15	0.458	0.623	0.025
29-Jun-12	R.Federer	Ad	37	25	62	28	0.187	0.266	0.755
29-Jun-12	R.Federer	Deuce	33	39	72	33	0.154	0.218	0.429
29-Jun-12	J.Benneteau	Ad	39	26	65	42	0.992*	0.996	0.264
29-Jun-12	J.Benneteau	Deuce	57	20	77	33	0.714	0.806	0.034
02-Jul-12	R.Federer	Ad	26	22	48	24	0.347	0.46	0.635
02-Jul-12	R.Federer	Deuce	23	29	52	25	0.27	0.371	0.103
02-Jul-12	X.Malisse	Ad	30	20	50	25	0.44	0.559	0.54
02-Jul-12	X.Malisse	Deuce	44	11	55	21	0.792	0.893	0.841
04-Jul-12	R.Federer	Ad	17	10	27	14	0.484	0.649	0.458

Continued on next page

Table A.4 – *Continued from previous page*

			Serves			Runs test			
Year	Server	Court	L	R	Total	Runs (r)	$F(r-1)$	$F(r)$	$U[F(r-1), F(r)]$
04-Jul-12	R.Federer	Deuce	10	19	29	13	0.25	0.399	0.283
04-Jul-12	M.Youzhny	Ad	15	16	31	20	0.865	0.929	0.48
04-Jul-12	M.Youzhny	Deuce	21	13	34	23	0.977*	0.991	0.803
04-Jul-12	D.Ferrer	Ad	33	41	74	52	0.999*	0.999	0.565
04-Jul-12	D.Ferrer	Deuce	24	31	55	31	0.75	0.829	0.299
04-Jul-12	A.Murray	Ad	38	27	65	30	0.214	0.297	0.876
04-Jul-12	A.Murray	Deuce	36	29	65	31	0.253	0.34	0.986
06-Jul-12	A.Murray	Ad	21	24	45	21	0.189	0.282	0.999
06-Jul-12	A.Murray	Deuce	30	24	54	26	0.273	0.372	0.806
06-Jul-12	J.Tsonga	Ad	42	17	59	21	0.065	0.116	0.23
06-Jul-12	J.Tsonga	Deuce	35	22	57	28	0.441	0.554	0.339
06-Jul-12	N.Djokovic	Ad	25	20	45	19	0.074	0.127	0.119
06-Jul-12	N.Djokovic	Deuce	30	22	52	29	0.728	0.814	0.433
06-Jul-12	R.Federer	Ad	21	20	41	18	0.103	0.172	0.64
06-Jul-12	R.Federer	Deuce	22	23	45	20	0.114	0.183	0.478
08-Jul-12	R.Federer	Ad	35	25	60	25	0.064	0.105	0.235
08-Jul-12	R.Federer	Deuce	22	33	55	25	0.205	0.294	0.348
08-Jul-12	A.Murray	Ad	52	21	73	32	0.566	0.675	0.094
08-Jul-12	A.Murray	Deuce	53	20	73	26	0.088	0.146	0.491
24-Jun-13	V.Hanescu	Ad	12	9	21	14	0.844	0.929	0.075
24-Jun-13	V.Hanescu	Deuce	13	11	24	14	0.596	0.747	0.784
24-Jun-13	R.Federer	Ad	9	11	20	9	0.132	0.257	0.153
24-Jun-13	R.Federer	Deuce	14	11	25	9	0.022	0.056	0.183
24-Jun-13	R.Nadal	Ad	29	17	46	22	0.382	0.508	0.088
24-Jun-13	R.Nadal	Deuce	40	10	50	21	0.943	0.978	0.289
24-Jun-13	S.Darcis	Ad	18	31	49	28	0.876	0.929	0.625
24-Jun-13	S.Darcis	Deuce	20	34	54	25	0.309	0.419	0.107
25-Jun-13	N.Djokovic	Ad	22	17	39	15	0.03	0.061	0.259
25-Jun-13	N.Djokovic	Deuce	19	21	40	19	0.215	0.32	0.172
25-Jun-13	F.Mayer	Ad	24	17	41	17	0.075	0.133	0.994
25-Jun-13	F.Mayer	Deuce	28	17	45	21	0.297	0.416	0.782
25-Jun-13	D.Ferrer	Ad	20	18	38	22	0.695	0.8	0.036
25-Jun-13	D.Ferrer	Deuce	11	16	27	18	0.92	0.965	0.647
25-Jun-13	M.Alund	Ad	31	15	46	22	0.538	0.668	0.117
25-Jun-13	M.Alund	Deuce	22	14	36	23	0.941	0.972	0.112
25-Jun-13	A.Ramos	Ad	19	6	25	13	0.912	0.972	0.317
25-Jun-13	A.Ramos	Deuce	25	14	39	17	0.193	0.304	0.236
25-Jun-13	J.Del Potro	Ad	8	25	33	17	0.950**	0.983	0.293
25-Jun-13	J.Del Potro	Deuce	18	12	30	14	0.23	0.363	0.12
26-Jun-13	J-W.Tsonga	Ad	22	13	35	15	0.147	0.248	0.211
26-Jun-13	J-W.Tsonga	Deuce	17	17	34	13	0.027	0.058	0.172
26-Jun-13	E.Gulbis	Ad	15	14	29	19	0.873	0.935	0.392
26-Jun-13	E.Gulbis	Deuce	13	18	31	16	0.411	0.56	0.201
26-Jun-13	F.Verdasco	Ad	28	13	41	19	0.462	0.607	0.502
26-Jun-13	F.Verdasco	Deuce	22	20	42	21	0.324	0.443	0.525
26-Jun-13	J.Benneteau	Ad	21	26	47	29	0.898	0.941	0.974
26-Jun-13	J.Benneteau	Deuce	27	25	52	35	0.982*	0.991	0.442
27-Jun-13	N.Djokovic	Ad	18	14	32	18	0.607	0.738	0.012
27-Jun-13	N.Djokovic	Deuce	17	20	37	17	0.166	0.264	0.267
27-Jun-13	B.Reynolds	Ad	16	22	38	22	0.747	0.842	0.766
27-Jun-13	B.Reynolds	Deuce	15	12	27	15	0.526	0.678	0.755
27-Jun-13	R.Gasquet	Ad	16	23	39	23	0.811	0.888	0.974
27-Jun-13	R.Gasquet	Deuce	28	20	48	23	0.29	0.401	0.216
27-Jun-13	G.Soeda	Ad	26	20	46	24	0.486	0.606	0.586
27-Jun-13	G.Soeda	Deuce	26	22	48	30	0.914	0.952	0.301
27-Jun-13	J.Levine	Ad	25	16	41	24	0.839	0.907	0.164
27-Jun-13	J.Levine	Deuce	24	16	40	22	0.667	0.778	0.592
27-Jun-13	J.Del Potro	Ad	7	29	36	11	0.164	0.334	0.759
27-Jun-13	J.Del Potro	Deuce	20	9	29	13	0.342	0.515	0.021
28-Jun-13	J.Melzer	Ad	31	11	42	15	0.132	0.239	0.377
28-Jun-13	J.Melzer	Deuce	24	20	44	23	0.461	0.583	0.408
28-Jun-13	S.Stakhovsky	Ad	17	13	30	11	0.023	0.054	0.583
28-Jun-13	S.Stakhovsky	Deuce	17	19	36	18	0.312	0.44	0.117
01-Jul-13	J.Janowicz	Ad	27	29	56	26	0.174	0.252	0.125
01-Jul-13	J.Janowicz	Deuce	30	29	59	35	0.853	0.905	0.151
01-Jul-13	J.Melzer	Ad	44	14	58	24	0.676	0.794	0.985
01-Jul-13	J.Melzer	Deuce	27	24	51	23	0.133	0.204	0.899
01-Jul-13	M.Youzhny	Ad	27	13	40	19	0.492	0.636	0.601
01-Jul-13	M.Youzhny	Deuce	22	18	40	23	0.708	0.808	0.929
01-Jul-13	A.Murray	Ad	28	14	42	21	0.615	0.741	0.118
01-Jul-13	A.Murray	Deuce	35	13	48	18	0.18	0.293	0.98
03-Jul-13	N.Djokovic	Ad	20	20	40	21	0.436	0.563	0.187
03-Jul-13	N.Djokovic	Deuce	25	13	38	17	0.278	0.412	0.351
03-Jul-13	T.Berdych	Ad	22	12	34	16	0.346	0.495	0.699
03-Jul-13	T.Berdych	Deuce	15	22	37	21	0.717	0.821	0.816
03-Jul-13	F.Verdasco	Ad	45	21	66	34	0.865	0.918	0.523
03-Jul-13	F.Verdasco	Deuce	29	34	63	32	0.418	0.52	0.001
03-Jul-13	A.Murray	Ad	9	39	48	17	0.664	0.818	0.449
03-Jul-13	A.Murray	Deuce	32	30	62	30	0.263	0.353	0.063

Continued on next page

Table A.4 – *Continued from previous page*

Year	Server	Court	Serves L	Serves R	Serves Total	Runs test: Runs (r)	$F(r-1)$	$F(r)$	$U[F(r-1), F(r)]$
05-Jul-13	N.Djokovic	Ad	47	27	74	28	0.024	0.042**	0.948
05-Jul-13	N.Djokovic	Deuce	60	24	84	30	0.059	0.098	0.012
05-Jul-13	J.Del Potro	Ad	32	46	78	33	0.07	0.108	0.568
05-Jul-13	J.Del Potro	Deuce	45	15	60	19	0.04	0.081	0.027
05-Jul-13	J.Janowicz	Ad	33	22	55	27	0.399	0.511	0.175
05-Jul-13	J.Janowicz	Deuce	32	24	56	27	0.297	0.399	0.84
05-Jul-13	A.Murray	Ad	25	22	47	20	0.073	0.123	0.205
05-Jul-13	A.Murray	Deuce	30	20	50	20	0.05	0.09	0.266
07-Jul-13	N.Djokovic	Ad	28	18	46	20	0.142	0.224	0.77
07-Jul-13	N.Djokovic	Deuce	24	19	43	22	0.412	0.536	0.537
07-Jul-13	A.Murray	Ad	27	18	45	23	0.487	0.611	0.981
07-Jul-13	A.Murray	Deuce	27	19	46	24	0.524	0.643	0.967
23-Jun-14	G.Dimitrov	Ad	17	17	34	14	0.058	0.111	0.042
23-Jun-14	G.Dimitrov	Deuce	19	17	36	24	0.938	0.97	0.298
23-Jun-14	R.Harrison	Ad	19	17	36	18	0.312	0.44	0.128
23-Jun-14	R.Harrison	Deuce	20	17	37	19	0.384	0.516	0.351
23-Jun-14	A.Murray	Ad	15	21	36	17	0.243	0.363	0.748
23-Jun-14	A.Murray	Deuce	27	12	39	16	0.209	0.334	0.959
23-Jun-14	D.Goffin	Ad	22	14	36	17	0.282	0.413	0.589
23-Jun-14	D.Goffin	Deuce	25	18	43	18	0.079	0.138	0.95
24-Jun-14	P.Lorenzi	Ad	40	6	46	9	0.023	0.095	0.861
24-Jun-14	P.Lorenzi	Deuce	32	7	39	13	0.502	0.715	0.668
24-Jun-14	R.Federer	Ad	14	15	29	12	0.065	0.129	0.191
24-Jun-14	R.Federer	Deuce	16	14	30	13	0.099	0.181	0.039
24-Jun-14	S.Wawrinka	Ad	17	16	33	15	0.145	0.241	0.387
24-Jun-14	S.Wawrinka	Deuce	25	10	35	13	0.119	0.224	0.978
24-Jun-14	J.Sousa	Ad	26	6	32	10	0.225	0.44	0.91
24-Jun-14	J.Sousa	Deuce	12	21	33	14	0.144	0.248	0.386
25-Jun-14	G.Dimitrov	Ad	20	9	29	18	0.965**	0.988	0.948
25-Jun-14	G.Dimitrov	Deuce	16	17	33	15	0.145	0.241	0.267
25-Jun-14	L.Saville	Ad	26	7	33	12	0.387	0.599	0.965
25-Jun-14	L.Saville	Deuce	24	11	35	15	0.262	0.407	0.272
25-Jun-14	A.Murray	Ad	8	18	26	11	0.227	0.392	0.226
25-Jun-14	A.Murray	Deuce	10	15	25	12	0.261	0.415	0.859
25-Jun-14	B.Rola	Ad	14	3	17	6	0.344	0.693	0.774
25-Jun-14	B.Rola	Deuce	15	8	23	9	0.082	0.18	0.142
26-Jun-14	G.Muller	Ad	20	14	34	15	0.142	0.239	0.593
26-Jun-14	G.Muller	Deuce	30	12	42	19	0.554	0.699	0.768
26-Jun-14	R.Federer	Ad	19	13	32	18	0.654	0.779	0.622
26-Jun-14	R.Federer	Deuce	16	16	32	11	0.009	0.024*	0.216
26-Jun-14	L.Rosol	Ad	19	29	48	20	0.086	0.145	0.451
26-Jun-14	L.Rosol	Deuce	17	34	51	20	0.091	0.156	0.32
26-Jun-14	R.Nadal	Ad	31	14	45	27	0.985*	0.994	0.4
26-Jun-14	R.Nadal	Deuce	36	13	49	24	0.897	0.949	0.748
26-Jun-14	S.Wawrinka	Ad	35	24	59	30	0.502	0.609	0.897
26-Jun-14	S.Wawrinka	Deuce	39	29	68	40	0.904	0.94	0.519
26-Jun-14	Y-H.Lu	Ad	32	24	56	24	0.087	0.139	0.803
26-Jun-14	Y-H.Lu	Deuce	20	33	53	33	0.974**	0.987	0.33
27-Jun-14	N.Djokovic	Ad	20	13	33	15	0.201	0.32	0.666
27-Jun-14	N.Djokovic	Deuce	17	17	34	18	0.43	0.569	0.254
27-Jun-14	G.Simon	Ad	23	23	46	25	0.559	0.672	0.796
27-Jun-14	G.Simon	Deuce	20	22	42	26	0.866	0.922	0.558
28-Jun-14	S.Giraldo	Ad	24	13	37	15	0.108	0.192	0.879
28-Jun-14	S.Giraldo	Deuce	21	10	31	17	0.793	0.892	0.384
28-Jun-14	R.Federer	Ad	19	14	33	15	0.171	0.278	0.489
28-Jun-14	R.Federer	Deuce	19	15	34	16	0.211	0.327	0.942
28-Jun-14	M.Kukushkin	Ad	23	33	56	31	0.747	0.827	0.873
28-Jun-14	M.Kukushkin	Deuce	22	32	54	27	0.435	0.548	0.678
28-Jun-14	R.Nadal	Ad	12	16	28	20	0.97**	0.988	0.317
28-Jun-14	R.Nadal	Deuce	29	9	38	19	0.958**	0.985	0.488
30-Jun-14	S.Wawrinka	Ad	18	11	29	20	0.974**	0.99	0.212
30-Jun-14	S.Wawrinka	Deuce	20	9	29	11	0.097	0.197	0.057
30-Jun-14	D.Istomin	Ad	13	14	27	17	0.786	0.882	0.888
30-Jun-14	D.Istomin	Deuce	19	13	32	19	0.779	0.873	0.475
02-Jul-14	A.Murray	Ad	18	16	34	18	0.438	0.577	0.563
02-Jul-14	A.Murray	Deuce	20	15	35	18	0.41	0.549	0.404
02-Jul-14	G.Dimitrov	Ad	17	21	38	26	0.971**	0.987	0.923
02-Jul-14	G.Dimitrov	Deuce	17	20	37	19	0.384	0.516	0.986
02-Jul-14	S.Wawrinka	Ad	34	21	55	19	0.007	0.015*	0.029
02-Jul-14	S.Wawrinka	Deuce	34	26	60	28	0.215	0.3	0.061
02-Jul-14	R.Federer	Ad	35	20	55	29	0.726	0.815	0.094
02-Jul-14	R.Federer	Deuce	33	19	52	27	0.662	0.764	0.063
04-Jul-14	N.Djokovic	Ad	35	27	62	28	0.149	0.218	0.435
04-Jul-14	N.Djokovic	Deuce	25	37	62	29	0.266	0.36	0.381
04-Jul-14	G.Dimitrov	Ad	23	29	52	28	0.594	0.699	0.112
04-Jul-14	G.Dimitrov	Deuce	29	19	48	27	0.781	0.86	0.074
04-Jul-14	R.Federer	Ad	19	14	33	17	0.41	0.554	0.221
04-Jul-14	R.Federer	Deuce	18	18	36	10	0	0.002*	0.533
04-Jul-14	M.Raonic	Ad	27	10	37	15	0.32	0.483	0.516

Continued on next page

Table A.4 – *Continued from previous page*

Year	Server	Court	Serves			Runs test			
			L	R	Total	Runs (r)	$F(r-1)$	$F(r)$	$U[F(r-1), F(r)]$
04-Jul-14	M.Raonic	Deuce	15	19	34	13	0.031	0.065	0.546
06-Jul-14	N.Djokovic	Ad	35	29	64	39	0.929	0.957	0.22
06-Jul-14	N.Djokovic	Deuce	47	28	75	27	0.008	0.016*	0.527
06-Jul-14	R.Federer	Ad	40	36	76	40	0.555	0.644	0.04
06-Jul-14	R.Federer	Deuce	35	52	87	34	0.018	0.030**	0.399
Sum			8,145	6,513	14,658	7,004	-	-	-
Mean			25.5	20.4	45.8	21.9	-	-	-

References

Binmore, K., Swierzbinski, J., & Proulx, C. (2001). Does minimax work? An experimental study. *The Economic Journal*, *111*(473), 445–464.

Chiappori, P.-A., Levitt, S., & Groseclose, T. (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *The American Economic Review*, 1138–1151.

Cross, R., & Pollard, G. (2009). Grand Slam men's singles tennis 1991–2009 serve speeds and other related data. *Coaching & Sport Science Review*.

Davey, P. R., Thorpe, R. D., & Williams, C. (2002). Fatigue decreases skilled tennis performance. *Journal of Sports Sciences*, *20*(4), 311–318.

Dohmen, T., & Sonnabend, H. (2016). Further field evidence for minimax play. *Journal of Sports Economics*, 1–18.

Gibbons, J. D., & Chakraborti, S. (2011). *Nonparametric statistical inference*. Springer.

Gilbert, B., & Jamison, S. (2013). *Winning ugly: Mental warfare in tennis–lessons from a master*. Simon and Schuster.

González-Díaz, J., Gossner, O., & Rogers, B. W. (2012). Performing best when it matters most: Evidence from professional tennis. *Journal of Economic Behavior & Organization*, *84*(3), 767–781.

Hsu, S.-H., Huang, C.-Y., & Tang, C.-T. (2007). Minimax play at Wimbledon: Comment. *The American Economic Review*, *97*(1), 517–523.

Klaassen, F., & Magnus, J. R. (2014). *Analyzing Wimbledon: The power of statistics*. Oxford University Press, USA.

Kovash, K., & Levitt, S. D. (2009). *Professionals do not play minimax: Evidence from Major League Baseball and the National Football League*.

Levitt, S. D., List, J. A., & Reiley, D. H. (2010). What happens in the field stays in the field: Exploring whether professionals play minimax in laboratory experiments. *Econometrica*, *78*(4), 1413–1434.

Mood, A. M., Graybill, F., & Boes, D. C. (1974). *Introduction to the theory of statistics*. McGraw-Hill, New York.

Mookherjee, D., & Sopher, B. (1994). Learning behavior in an experimental matching pennies game. *Games and Economic Behavior*, *7*(1), 62–91.

Mookherjee, D., & Sopher, B. (1997). Learning and decision costs in experimental constant sum games. *Games and Economic Behavior*, *19*(1), 97–132.

O'Donoghue, P. G. (2001). The most important points in Grand Slam singles tennis. *Research Quarterly for Exercise and Sport*, *72*(2), 125–131.

O'Neill, B. (1987). Nonmetric test of the minimax theory of two-person zerosum games. *Proceedings of the National Academy of Sciences*, *84*(7), 2106–2109.

Palacios-Huerta, I. (2003). Professionals play minimax. *The Review of Economic Studies*, *70*(2), 395–415.

Palacios-Huerta, I. (2014). *Beautiful game theory: How soccer can help economics.* Princeton University Press.

Palacios-Huerta, I., & Volij, O. (2008). Experientia docet: Professionals play minimax in laboratory experiments. *Econometrica*, *76*(1), 71–115.

Rapoport, A., & Boebel, R. B. (1992). Mixed strategies in strictly competitive games: A further test of the minimax hypothesis. *Games and Economic Behavior*, *4*(2), 261–283.

Rosenthal, R. W., Shachat, J., & Walker, M. (2003). Hide and seek in Arizona. *International Journal of Game Theory*, *32*(2), 273–293.

Shachat. (2002). Mixed strategy play and the minimax hypothesis. *Journal of Economic Theory*, *104*(1), 189–226.

Wald, A., & Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, *11*(2), 147–162.

Walker, M., & Wooders, J. (2001). Minimax play at Wimbledon. *The American Economic Review*, *91*(5), 1521–1538.

Wolfers, J., & Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities.* (Working paper)

Wooders, J. (2010). Does experience teach? Professionals and minimax play in the lab. *Econometrica*, *78*(3), 1143–1154.

A.3

What is the effect of insider trading on price efficiency? Evidence from a betting exchange

Paolo Bizzozero, Raphael Flepp, Egon Franck

Abstract

We present evidence that insider trading substantially contributes to the price discovery process after important news events and thus helps to create efficient markets. Live betting offers a unique opportunity to isolate and measure the activity of traders with earlier access to information (insiders). We perform an event study using detailed, point-by-point data from 141 men's singles matches at two major professional tennis tournaments. The results show that betting prices start updating long before the general public receives the new information, indicating the existence of insider traders. Most importantly, the cumulative abnormal return during the first few seconds of insider trading following an important event is more than 60% of the full price reaction observed once the public receives the new information, meaning that insider trading has a large impact on price discovery. We also show that a simple trading strategy based on inside information can generate significant returns.

JEL Classification: G14, L83

Keywords: Market efficiency, insider trading, event study, tennis betting

[†]A version of this paper has been submitted to *Journal of Economic Behavior & Organization.*

1 Introduction

In theory, the extent to which insider trading affects the efficiency of security market prices is uncertain. One view proposes that insiders, i.e., traders with monopolistic access to any pricing-relevant information, contribute to more efficient markets by quickly and accurately incorporating new information into prices (Manne, 1966; Carlton & Fischel, 1983). Another view proposes that the existence of insiders can cause outsiders not to trade: as outsiders will also stop collecting information, the market will be less informationally efficient (Fishman & Hagerty, 1992). Depending on which effect dominates, insider trading can have an overall positive or negative effect on the efficiency of security prices.

In the corporate world, insider trading is the practice of trading securities by corporate insiders, like top officers and directors, who have privileged access to information about the true value of their firm's stock. Most financial regulators further distinguish between legal insider trading, which is based on non-material (public or nonpublic) information, and illegal insider trading, which is based on material nonpublic information. Although no statutory definition of "materiality" is available, material information is generally defined as information expected to significantly affect the stock price of a company, like major acquisitions or changes in top management.

Despite a rich literature on price efficiency, direct empirical evidence on whether and to what extent insider trading based on material information affects the efficiency of security market prices is scarce due to its illegality in most countries and the consequent lack of data. The vast majority of studies related to the efficiency of stock prices show that (1) self-reported (legal) insider trades correctly predict future stock performance, meaning that corporate insiders can gain abnormal profits (for a review, see Lakonishok & Lee, 2001) and that (2) stock prices react fully and swiftly after corporate news announcements.[1] However, these studies do not investigate *material*

[1]For example, see Ikenberry, Lakonishok, and Vermaelen (1995) for stock splits announcements, Mitchell, Pulvino, and Stafford (2004) for mergers announcements, and MacKinlay (1997) for earnings announcements.

insider trades. The only relevant papers on the subject are those of Cornell and Sirri (1992), Meulbroek (1992), and Chakravarty and McConnell (1997), which analyze the market's reaction to illegal insider trades (the data originate from criminal and civil litigation reports) and show that insider trades lead to more rapid price discovery.[2] However, Chakravarty and McConnell (1999) disputed these studies by showing, using a refined methodology to re-analyze the data of those previous studies, that the effect of insider trades on prices is not discernibly different from that of non-insider trades. Thus, there is generally little and mixed evidence on the topic.

In this article, we present an analysis of unique and naturally occurring field data that provide a rare opportunity to isolate the effect of insider trading on securities price efficiency: live tennis betting. In this context, most betting traders follow the match on TV or via internet score feeds and, whenever new relevant information becomes observable, they update their bets online.[3] Our analysis takes advantage of the inevitable technical delays in the transmission of match information from the stadium to end receivers. For example, the delay from filming to receiving a TV broadcast is *at least* five seconds (Kooij, Stokking, van Brandenburg, & de Boer, 2014; Hutchins, 2014; Brown & Yang, 2017). Due to this delay, the insider traders sitting in the stadium benefit from a fleeting informational advantage compared to "live"-TV, "live"-scoreboard, and other slower traders because insiders observe important information before them. Thus, we can attribute the price update in the first five seconds after important news events during the match to the activity of insider traders.

More specifically, we analyze the price reaction after a player wins a *set*.[4] Winning a set is an important step toward victory: therefore, observing in advance which player won the set constitutes a considerable informational advantage to insiders. We use high-frequency betting data from Betfair, one of the largest betting exchanges

[2]Using U.S. data, Cornell and Sirri (1992) and Chakravarty and McConnell (1997) analyze two specific cases of insider trading in the stock of two takeover target firms, whereas Meulbroek (1992) analyzes 183 insider trading cases from different firms.

[3]We use trader as a synonym for a bettor, gambler, or wagerer.

[4]The Appendix A.1 provides a short introduction to tennis scoring rules and terminology.

worldwide, on 141 tennis men's singles matches played at two major professional tennis tournaments, the French Open and the Wimbledon Championships, over the 2009–2014 period. We complement this information with match data about players, courts, winner, start and end match times. Our dataset provides us with two practical advantages with respect to standard financial data. First and most important, our second-by-second data allow us to determine the event time (when new information is released) with high accuracy and thus to measure with precision the impact of insider trading activity on prices. Second, because the players take a short break after the end of each set (a low-information period), we do not have to deal with the problem of confounding events.

Using event study methods, we find that the cumulative abnormal returns averaged across the 365 event observations ($CAAR$) start increasing immediately after the set events: the $CAAR$ values at one second and five seconds following the event are 0.82% and 3.06%, respectively, and are significant at the 1% level. Most importantly, we show that the cumulative abnormal return during the first five insider trading seconds following the set events—when the TV viewers and other slower traders have not yet seen the event—accounts for more than 60% of the full price reaction observed once the public receives the new information. Furthermore, we show that the impact of insider trading is even larger for unanticipated news events, like *tie-break set* events, when inside information is more valuable. Finally, we estimate that a simple dynamic back-lay trading strategy implemented in the seconds after the event can yield large, risk-free profits to insiders, varying between 5% and 7%.

In this article, we can uncover the impact of insider trading in a real-world environment. To the best of our knowledge, this paper is the first to provide a clear setting for testing the question as to how insider trading influences the efficiency of price discovery. Overall, we provide important evidence that insider traders significantly contribute to higher price efficiency and thus market quality.

The remainder of this article is organized as follows. Section 2 reviews the applicable literature. Section 3 describes our setting. Section 4 formulates the hypotheses and presents the data. Section 5 outlines our empirical methodology. Section 6 presents the results. Section 7 discusses the results and concludes.

2 Literature review

The debate about the pros and cons of insider trading is important as it might influence the decision on whether and to what extent to regulate such trading in financial markets (Leland, 1992). Researchers and regulators have examined both the fairness and the economic implications of insider trading. Concerning the latter, which is the focus of our article, the key issue involves the assessment of the impact of insider trading on price efficiency. In this regard, two contradictory views have emerged. The first view, pioneered by the work of Manne (1966) and extended by Carlton and Fischel (1983), claims that trading on inside information leads to more informationally efficient stock prices. The main argument of the supporters of insider trading is that insider trades are informative and convey precious information about the future performance of a firm and thus about its true value (Leland, 1992). Accurate stock prices help the capital market to efficiently allocate its resources: for example, a takeover decision is often based on a stock-price-based estimate of the target's value.

The second view claims that insider trading leads to less informationally efficient stock prices due to adverse selection costs (Fishman & Hagerty, 1992). In this logic, some outsiders perceive the market to be unfair and thus might stop trading and searching for relevant information because they anticipate that insiders still have a superior knowledge. As a consequence, prices will be less informationally efficient and more volatile, and market liquidity will decrease (Leland, 1992).

According to Fama (1970), asset prices are efficient if they fully reflect all available information, public or private. Several financial studies have tested the semi-strong form of the EMH, which states that security prices include all relevant public information swiftly and fully, using event study methods. The evidence is mostly supportive because stock prices adjust rapidly, usually within a day, to announcements (for an overview, see Fama, 1991).[5] In general, event studies using stock returns face a number of issues, including the difficulty in determining the precise event date (the moment information is publicly disseminated), the effect of confounding events around the event date, and the need to use an asset pricing model to evaluate the normal stock returns.[6]

Other studies test the strong-form of the EMH, which states that prices include all public and private information, by investigating whether corporate insiders can gain abnormal profits when trading their company' stock (Lakonishok & Lee, 2001). Rozeff and Zaman (1988), for example, find evidence that managers profitably trade their companies' shares before important corporate events and gain abnormal profits, suggesting that their trades are based on important information. However, these articles mostly investigate *legal* insider trading, i.e., trading on *non*-material inside information that is self-reported by corporate insiders to the local financial authorities.[7] A drawback of these studies lies in their inherent inability to precisely determine the motivation behind these trades: for example, managers might trade their firm's stock for hedging, risk-sharing, or liquidity reasons (Jaffe, 1974).

Most relevant to price efficiency studies is *illegal* insider trading activity, i.e., trading on *material* inside information. However, corporate insiders obviously refrain from reporting violations of the law to authorities, so the real amount of insider trading activity cannot be determined with precision (Keown & Pinkerton, 1981).

[5] Some researchers have also observed a pre-announcement drift in stock prices (a so-called price run-up) and suggest that this drift is caused by illegal insider trading on leaked material information (Mitchell et al., 2004).

[6] This latter issue is known as the joint hypothesis problem. Stock efficiency studies rely on some equilibrium asset pricing models to measure any abnormal return, i.e., the expected return given the absence of a particular event like an earnings announcement. Thus, any abnormal return observed may reflect a market inefficiency, an inaccurate pricing model, or both (Fama, 1991).

[7] Corporate insiders must disclose their trades in their firms' securities to the local financial market authorities. In the U.S., the Securities and Exchange Commission (SEC) publishes monthly a list of all self-reported insider trades in the *Official Summary of Insider Transactions*. Most studies are based on these SEC filings. However, the filings are too vague to characterize the type of informational advantage held by managers (Ke, Huddart, & Petroni, 2003).

Thus, despite the importance of the subject, the empirical evidence on the impacts of trading on material information on price efficiency is scarce. One important exception is the study of Meulbroek (1992), who uses information on illegal insider trading from 183 civil cases brought by the Securities and Exchange Commission (SEC) in order to investigate the impact of these informed insider trades on the efficiency of stock prices. She finds that insider trading cause significant movements in the prices, thereby increasing their accuracy. Notably, Meulbroek shows that between 40% and 50% of the price adjustment following the release of inside information to the public is caused by insider trading.

Two other studies, Cornell and Sirri (1992) and Chakravarty and McConnell (1997), also show that insider trading leads to more efficient stock prices by analyzing illegal insider trades in the stock of two acquisition targets, Campbell Taggart and Carnation, respectively. However, Chakravarty and McConnell (1999) dispute these three previous studies on the grounds that they neglected to distinguish the effect of insider trading from that of non-insider trading on price discovery. Using a refined methodology, Chakravarty and McConnell re-analyze the data of those previous studies and show that the effect of insider trades on prices is not discernibly different from that of non-insider trades.

Some authors have used the data from the betting markets to investigate the semi-strong form of the EMH of betting prices during sports events. For example, Croxson and Reade (2014) investigate the reaction of prices on a betting exchange to soccer goals scored within five minutes before the half-time break. Since during half-time there is little new information, one can test if the prices quickly adjust to the news of the goal and stay constant thereafter (no price drift). The authors conclude that the prices update swiftly and fully and that the betting market is semi-strong form efficient. In contrast, Choi and Hui (2014) reject the semi-strong form of the EMH using very similar in-play soccer data because prices overreact after surprising goals.

Most related to our study is the work of Schnytzer and Shilony (1995), who also employ a setting in which one can reliably distinguish between the activity of two groups, one with and one without access to inside information. Schnytzer and Shilony use the horse pari-mutuel betting setting in Melbourne.[8] Pari-mutuel betting is often offered at the racetrack next to over one hundred standard bookmakers and also at off-track facilities. In this setting, pari-mutuel traders at the racetrack have the advantage of observing how the bookmaker odds change shortly before the race. For example, they may observe that the bookmaker odds on a given horse have significantly decreased, suggesting that many traders, some of whom may have held inside information, favor that horse. Therefore, insiders profit from having a source of "second-hand" inside information via the bookmakers' odds (Schnytzer & Shilony, 1995). In contrast, as communications to outside the stadium are prohibited by law, off-track pari-mutuel traders are not informed about the last-minute developments and thus have an informational disadvantage. By analyzing the bets of these two isolated groups, Schnytzer and Shilony find that insiders make larger profits and that inside information on average correctly predicts the outcome of the race. However, their paper does not investigate the relationship between insider trading and price efficiency.

Finally, Brown (2012) analyzes the bid-ask spreads of the bookmakers' odds during the 2009 Wimbledon men's tennis final and asks whether some traders have an advantage due to superior analytical skills or due to access to material inside information. He observes an increase in the in-play bid-ask spreads prior to and during public information arrival and attributes this to an increase in asymmetric information. Brown proposes that, in the presence of insiders, the bookmakers increase the bid-ask spread to offset losses to the informed traders. However, Brown provides no direct evidence on the effect of insider trading on price efficiency.

[8]The pari-mutuel, or totalisator, is a betting system in which all the bets on the horses are pulled together in a so-called pool. The racetrack organizers handle the process by pooling and distributing the money, thereby earning a fee. The odds on different horses vary until the pool is closed prior to the start of the event; at that moment, the payoff odds are determined.

Overall, the empirical evidence on the economic effects of insider trading on market efficiency is either little or mixed in both the classical financial markets and the betting markets. This article contributes to closing this literature gap by providing a simpler setting that allows us to overcome the difficulties of previous investigations.

3 Background information

To understand the analyses in this paper, we introduce the advantages of using a betting exchange as a research laboratory and provide some important background information on in-play tennis betting. Most importantly, we discuss how we can distinguish between the trading activity of insider and outsider traders in our setting.

Compared to the research in standard financial markets, which mostly focuses on stock prices, research in betting markets offers several advantages. First, whereas equity is infinitely lived, a betting asset (a bet) is short-lived because the fundamental value of a bet is revealed at contract expiration. Therefore, one can test how efficiently the prices of betting assets (the odds) predict an outcome, like the winner of a match. Second, compared to stock prices, betting odds have two attractive properties: (1) new information has a larger impact on a bet value because a bet has a high probability of default (Brown, 2012); (2) bet trading, especially during the match, is mainly motivated by information about the fundamental value of a bet (Williams, 2005), whereas stock trading is motivated by a multitude of factors beyond the information about the fundamental value of equity, e.g., portfolio rebalancing or liquidity needs.[9]

Betting exchanges have several interesting features. Different from bookmakers, exchanges provide an online market for opinions where traders bet against other traders (without an intermediary) by offering and accepting odds under which they are willing to buy (*back*) or sell (*lay*) a certain bet. Thus, a betting exchange works very similarly to a classical stock exchange. In this article, we use data from one

[9]Uninformed or sentimental trading may occur in both markets.

of the major online betting exchanges worldwide, Betfair. The trading volumes on Betfair are very large: for example, 1.2 billion bets were placed on the exchange in 2014, corresponding to three million bets per day, resulting in a total matched volume of $92 billion (Betfair, 2015). In the United Kingdom, the total revenue generated by tennis betting is second only to that generated by soccer betting (Townend, 2016).

Betfair revolutionized the sports betting industry in 2000 when it introduced the in-play (or live) feature, allowing traders to continuously place or accept bets as the competition unfolds. Tennis in-play betting is very popular for several reasons. First, trading in tennis is simpler than in other sports because there are just two contingencies for each point: a player wins or loses it. Traders mostly follow the match on TV, and whenever new information becomes observable, they update their bets or place new ones (Brown & Yang, 2017). Second, liquidity, which is indispensable for trading, is generally high. Third, points are scored frequently, and in the space of seconds, a match can swing from a match point for one player to a match point for the other player; thus, the constant fluctuation in the odds facilitates trading. Finally, there are many betting opportunities because there are many tournaments throughout the year.

In-play betting constitutes a predominant share of the total betting volume generated in tennis. In our sample of 141 matches, $2.5 billion worth of bets were placed in-play, corresponding to 85% of the total (pre-play and in-play) volume. The maximum in-play volume in our sample, $70.5 million, was reached during the 2013 Wimbledon men's semifinal between Djokovic and Del Potro.

In the following two paragraphs, we briefly illustrate the basics of in-play betting. Fig. 1 depicts a screenshot of the in-play order book during the match between Murray and Wawrinka taken from *betfair.com* on 9th June, 2017.[10] The order book presents the bets on the match winner available in the order book: these are the bets

[10]This match, played at the French Open, serves only as illustration and is not included in our sample.

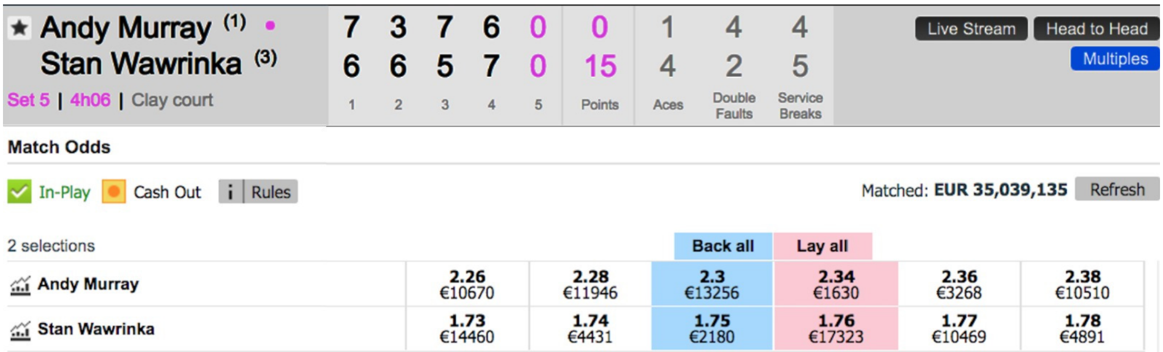


Figure 1
Displayed is a screenshot from *betfair.com* illustrating its interface. The screenshot was taken on 9th June, 2017 during the match between Murray and Wawrinka at the French Open tournament. Four hours into the game, the total matched amount was €35 million (approximately $39 million).

previously placed by other traders and that have not yet been matched. The third (fourth) column shows the current best back (lay) odds for each player. The available volumes (in €) are provided below the odds. According to Fig. 1, the best odds to bet on Murray are 2.3, and it is possible to bet up to €13,256 at this price. As the inverse of the match odds can be interpreted as the player's probability of winning the match, after four hours of play, Wawrinka was perceived as slight favorite for the victory, with a probability of roughly 57%.

The simplest way to bet on Murray is to click on 2.3 (under "Back all") and enter the desired stake: by doing so, the trader (the backer) enters into a contract with one or more traders (the layers) who are taking the opposite position (against Murray) and offering odds of 2.3.[11] The backer is placing a *market* order, thereby matching the outstanding orders previously placed by the layers. Assuming that the backer puts €10 at stake, the corresponding payoff to the backer if Murray wins the match is €13 plus the initial €10 stake.[12] As betting is a zero-sum game, the layer loses €13. If the backer wishes to back Murray at the higher odds of 2.38, the backer places a *limit* order, thereby offering odds higher than the current available market odds. If entered,

[11] Laying is an exclusive feature of betting exchanges in which layers are practically taking the role of a bookmakers offering the odds.

[12] We ignore the 5% commission on winning bets collected by Betfair (no commission is paid when a loss is incurred).

this order shows up on the "Lay all" side of the order book and can be matched by a trader willing to lay Murray at those odds.[13]

In tennis betting, a trader who can obtain information about the match earlier than others is an insider. Match information about the score or the physical and mental status of the players is material because it will impact the prices. The fastest way for traders to access this information is to sit directly in the stadium, which also allows them to see special situations, such as the signs of injury or the calls for physiotherapy assistance. As the venue where tennis is played is called the court, insider traders are sometimes referred to as courtsiders.[14]

Brad Hutchins, a former tennis courtsider, wrote in his book: "If you're working on centre court at a Grand Slam [tennis tournament], there could be up to twenty other traders gambling on court. [...] When courts are competitive, speed becomes the key. [...] If everyone gambling at home is watching a television feed that is delayed by five seconds, you have just five seconds to take advantage. That's plenty of time to log a point" (Hutchins, 2014, p. 28). Trading speed is essential for in-play traders, who have large incentives to quickly update their bets to gain a profit after each point. To trade alghorithmically, some sophisticated courtsiders transmit live score data to remotely located computers, which then place new orders on the betting exchange (Dickson, 2015).

In contrast, off-stadium traders (outsiders) rely on communication technologies to stay informed about the progress of the match. The most comprehensive source of information for outsiders are TV broadcasts; they not only provide score information (as

[13] The terms market and limit orders are also common in the standard financial markets. For example, placing a market order corresponds to either a buy or sell order at the current market price, whereas placing a limit order corresponds to either a buy order at a price below the current market price or a a sell order at a price above the current market price.

[14] Analogously, different terms exist for other sports, like pitchsiding in cricket and soccer. Although trading from the stadium was generally legal under French and British laws for the 2009–2014 period covered by our sample, such activity generally violates the terms and conditions of the ticket purchase. For example, Article 24 of the ticket terms and conditions at Wimbledon (2017) reads: "Betting is prohibited in the Grounds at all times", and Article 19 specifies that: "The use of photographic equipment, mobile telephones, computers, tablets or other electronic devices, communication devices, audio-visual equipment or radios must not [...] supply or transmit data for the purposes of betting or gambling (or assisting for these purposes)."

internet scoreboards) but also key information like a player's fatigue level, confidence, or tactics. However, TV images are significantly delayed in comparison to the real match. Betfair clearly indicates that any "transmissions described as "live" by some broadcasters may actually be delayed" and that the "extent of any such delay may vary, depending on the set-up through which they are receiving pictures or data."[15] Due to the large number of factors determining the length of the TV delay, accurate estimates are difficult. Kooij et al. (2014) measure the delay in live TV broadcasts in the Netherlands. They estimate that a minimum delay of four seconds is introduced by the steps of encoding and modulation of the images, on top of which one should add between one and six seconds for the transmission.[16]

Because Kooij et al.'s minimal delay of five seconds corresponds with the TV delay mentioned by Brad Hutchins and by Betfair, and other information sources like internet scoreboards are also delayed, we use five seconds as the cutoff for distinguishing the insider trading period (within five seconds after the event) from the outsider trading period (from six seconds after the event onward). In other words, the communication delays allow us to differentiate insider from outsider trading activity: any price movement observed within five seconds following an event would reveal the presence of insider traders, if any.

In an effort to slow down insider traders, Betfair has put in place a bet processing delay, a so-called speed bump. The speed bump imposes a delay of five seconds

[15]Betfair also writes that: "Although the current score, time elapsed, video and other data provided on this site is sourced from "live" score feeds provided by third parties, you should be aware that this data may be subject to a time delay and/or be inaccurate. Please also be aware that other Betfair customers may have access to data that is faster and/or more accurate than the data shown on the Betfair site. If you rely on this data to place bets, you do so entirely at your own risk. Betfair provides this data AS IS with no warranty as to the accuracy, completeness or timeliness of such data and accepts no responsibility for any loss (direct or indirect) suffered by you as a result of your reliance on it."

[16]Concerning the transmission of TV images, the length of the delay can greatly vary, depending on the transmitting technologies (satellite is slower than terrestrial, while internet streaming is the slowest), the quality (high definition is slower than standard definition), the subscription (digital is generally slower than analog TV), and the broadcaster (which use different hardware). The delay increases with the geographical distance between the event, the local broadcaster, and the household (Kooij et al., 2014). Fig. 1 in Kooij et al. (2014) illustrates the various steps of a TV content delivery chain: live TV broadcasts differ from pre-recorded content (like a TV series) because they introduce additional steps (and delays) due to the filming and transmission to the broadcaster from the event location. Besides, as the traders on a betting exchange are dispersed all over the world and use different technologies, an attempt to determine the actual average delay for all the traders is unrealistic.

between the time in which a market order (a new bet) is submitted and the time in which it is logged on the exchange and is eventually matched with outstanding orders (Brown & Yang, 2017). This mechanism is designed to protect the slower traders by giving them time to cancel or reduce the size of their outstanding (unmatched) orders without any delay.[17] In practice, once the market trader has entered the stake and confirmed the order, Betfair begins a five second countdown, at the end of which the market order is logged on the exchange; if the counterparty has not yet canceled the outstanding bet (within five seconds), the two orders are matched; if the counterparty has already canceled it, no transaction will take place.

However, the length of the speed bump may be inappropriate for protecting slower traders. First, as Kooij et al. (2014) shows, the actual TV delay may be well longer than five seconds. Besides, the traders also need some time to react and update their bets. Second, Brown and Yang (2017) suggest that insider traders may circumvent the speed bump by placing two limit orders, one on each player, and then timely canceling the wrong position once the new score data come in. Since canceling a limit order occurs without delay, insider traders can still profit from their time advantage. Dan Dobson, a former insider trader, confirms this practice in an interview with the BBC: "We had an automated system whereby the point data would come in and then we would cancel any [outstanding] bets that we had in the market that we deemed were at the wrong price [...]. Then we would place bets straight back into the market that we deemed were now the correct price" (Cox, 2015). Clearly, such strategies (which we illustrate with an example below) reduce the protection offered by the speed bump to slower traders.

[17]However, changing the odds of a limit order is treated as placing a new market order and is subjected to the speed bump.

4 Research hypothesis and data

We focus on the speed of the price adjustment to new information, which is one key aspect of market efficiency. Thus, our main null hypothesis is that insider trading does not affect the efficiency of betting prices.

Hypothesis 1. The efficiency of price discovery is unaffected by insider trading.

The market efficiency theory (e.g., Manne, 1966; Carlton & Fischel, 1983) predicts that because insider trading is based on new, relevant information, insiders contribute to price discovery by rapidly incorporating information into market prices. Therefore, the alternative hypothesis is that insider trading is directly linked to the speed of incorporating new information into the prices of financial securities (Fernandes & Ferreira, 2009). We test our hypothesis using event study methods, which allow us to observe when and how the betting prices update following important informational events. Given the minimum communication delay of five seconds, we will reject the null hypothesis if we observe a significant price adjustment during the first five insider trading seconds following the event.

In the following, we lay out three broad scenarios regarding the hypothetical impact of insider trading on the efficiency of price discovery. Fig. 2 depicts the earliest time at which the cumulative abnormal returns start increasing following an important event.[18] The horizontal axis represents the time elapsed (in seconds) after the event, which takes place at time zero. In all scenarios, we assume a minimum informational delay (for the outsiders vs. the insiders) of five seconds.

Scenario 1 depicts a situation without insiders in which all traders are, for example, TV traders. The traders see the delayed TV images five seconds after the actual event ($time = 0$) and place their market bets. After a five second delay due to the

[18] In Section 6, we describe the event study methodology in detail. The cumulative abnormal returns represent the sum over a specified period after the event of the abnormal returns, i.e., the difference between the realized returns and the expected (normal) returns. To understand Fig. 2, one must understand that the price adjustment begins when the cumulative abnormal returns start to increase.

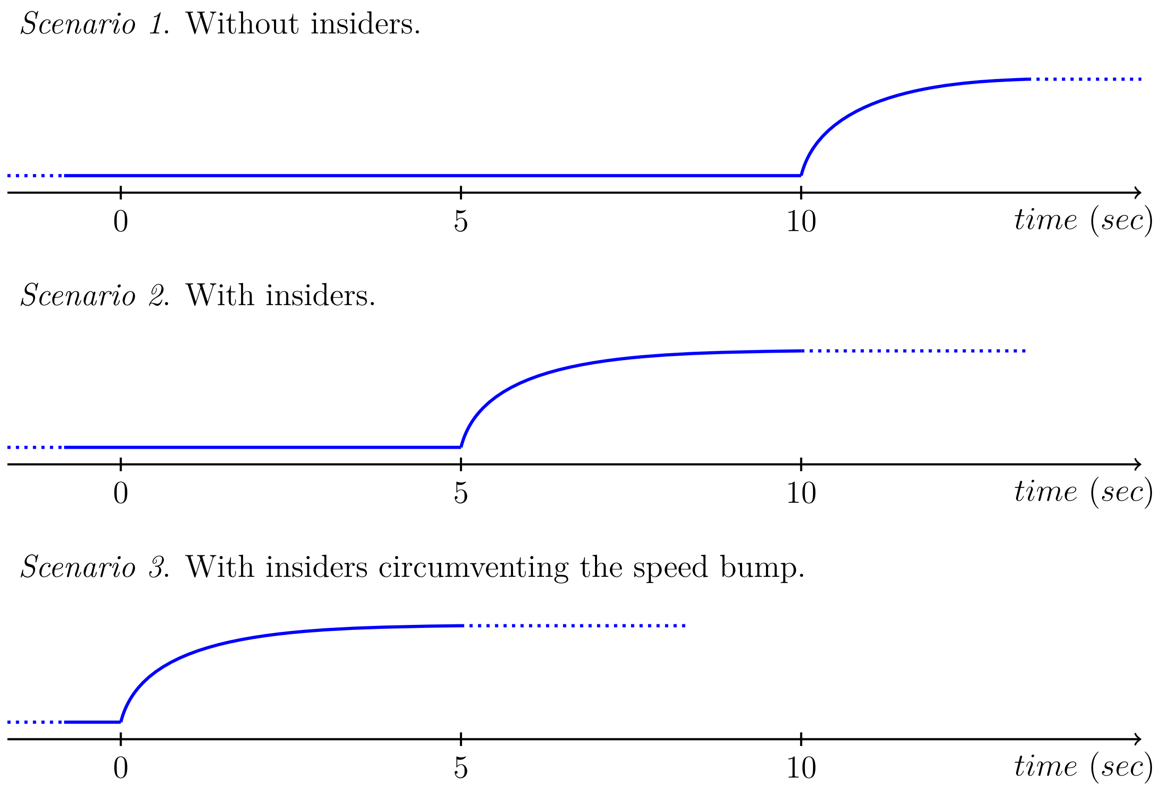


Figure 2
Displayed is the timeline of the event study under three scenarios. Time is in seconds, and zero corresponds to the second when a player wins a set. The line indicates the assumed time pattern of the cumulative average abnormal returns computed from the betting odds. The line should be flat until the moment at which some betting traders receive and act upon new material information, causing the betting odds to change.

speed bump, the new orders are logged on the exchange. Thus, the price adjustment begins, at the earliest, ten seconds following the event.

Scenario 2 depicts a situation with insiders and outsiders. If insiders observe the event as it happens ($time = 0$) and place new market orders that are subjected to the speed bump, the price adjustment begins, at the earliest, five seconds following the event.

Scenario 3 depicts the same situation as *Scenario 2*, with one exception: we assume that some sophisticated insider traders are able to circumvent the speed bump by timely canceling their limit orders on the wrong side of the market without any delay, as suggested by Brown and Yang (2017) and Dan Dobson. For example, insiders observe Federer winning the set and immediately cancel their orders for “Federer to lose” but not those on “Federer to win”. Contemporaneously, the slower traders, who are still unaware of the outcome of the set, will see their market orders on "Federer to win" matched but not those on "Federer to lose". Thus, under this scenario, the price adjustment begins immediately following the event. If the adjustment is significant, we will reject the null hypothesis.

Overall, the presence of insiders anticipates the start of the price adjustment by roughly five seconds (*Scenario 2*), which corresponds to the length of the informational delay. Furthermore, the presence of some sophisticated insiders further anticipates the price adjustment by roughly five seconds, because those traders avoid the speed bump (*Scenario 3*).

Our sample consists of 141 Grand Slam men's singles matches played between 2009 and 2014 at Roland Garros (61 matches) and at Wimbledon (80 matches).[19] The betting data originate from Betfair and are provided by Fracsoft, the official data vendor. For every second during a match (1,296,688 seconds in total), we have the best back and lay odds. The total matched volume is \$2.95 billion, 85% of which is generated in-play.

Detailed match data are provided by IBM, the official supplier of information technology to both tennis tournaments. Beyond general information about the match, such as players and the start and end match times, the IBM data also contain 31,018 point-level information on the score, time (exact to the second), and winner. The score is fed into the system directly by the match umpire using a computer, whereas other statistics are collected by analysts who attend the match and manually feed the data into the system.

Our final merged sample consists of 31,018 observations for 141 matches. The sample is heterogeneous, containing matches from the first stage up to the finals and a total of 83 unique players. Table 1 provides summary statistics. The average in-play matched volume is roughly \$18 million. Usually, semifinal and final matches attract

[19] Table A.1 in the Appendix provides the full list of matches.

Table 1
Summary statistics.

Variables	Mean	Std. dev.	Min.	Max.	N
Panel A: betting odds					
In-play matched volume (mio $)	18.1	19.4	0.2	70.5	1,296,688
In-play order processing delay (seconds)	5	0	5	5	1,296,688
Panel B: match information					
Duration (minutes)	154.7	50.1	69.9	284.8	141
Number of points	221.2	66.1	117	437	141
Number of sets	3.6	0.7	3	5	141
Number of games (in match)	35.8	10.1	20	77	141

Notes: The table reports the summary statistics for the 141 men's singles matches in our sample. The matches were played at the French Open and at the Wimbledon Championships between 2009 and 2014 (see Table A.1 in the Appendix). In Panel A, we report the statistics about the bedding odds from Betfair. In Panel B, we report the statistics about the match from IBM.

larger volumes than first-round matches, as shown by the $70.5 million matched during the 2013 Wimbledon semifinal between Djokovic and Del Potro. Different from other sports, the speed bump in tennis has remained fixed to five seconds over time. An average match lasts 155 minutes and consists of 3.6 sets, 36 games, and 221 points.

We use the end of a set—when a player wins a set—as the informational event, because it satisfies three important criteria. First, the end of a set is an important news event because it is a pivotal moment in a match. At the French Open and at Wimbledon, the player who first wins three sets wins the match: therefore, a player's probability of winning the match increases after winning a set. Second, the end of a set is an easily observable event. Third, a 120-second break follows the end of each set. This is advantageous because the event study results are less affected by overlapping and confounding events in the seconds following the end of a set.[20] Overall, our sample includes 365 set events. Since the betting market is immediately closed when a player wins the match, we exclude the last set from our analyses.[21]

In a further analysis, we analyze a subsample encompassing the sets won after a tie-break (N=79). A tie-break decides the outcome of very balanced sets, and its outcome is often unanticipated. When we compare the outcome of a tie-break with the outcome of a set won by a large margin (e.g., six games to zero), the tie-break

[20] Croxson and Reade (2014) considers the 15-minute halftime break in soccer matches as a "low-information period" because important information is rarely revealed during the break. Similarly, in tennis, little information is revealed during the break, except when players start showing clear signs of fatigue, stress, or pain.

[21] At the French Open and at Wimbledon, a match is played as the best of five sets. Since we do not include the last set when the match ends, we analyze between two and four set events per match. The "Number of sets" reported in Table 1 corresponds, however, to the statistics for the original sample before excluding the match points.

usually has a larger surprise component. In our sample, a tie-break lasts on average approximately eight minutes (with a minimum of four minutes and a maximum of 17 minutes), and thus the traders have enough time to recognize the importance of the moment and are prepared to adjust their bets.

Fig. 3 illustrates some characteristics of our two main events, the set and tie-break events, and provides insight into our data. More specifically, it shows the evolution of Del Potro's odds-implied winning probability (henceforth WP) over the match when he played against, and lost to, Djokovic at Wimbledon on 5th July, 2013.[22] As the match started, Del Potro's WP was 15.3%.

The vertical dashed lines indicate the end of a set (excluding the last one). Djokovic won the first set, resulting in a decrease of Del Potro's WP from 15.3% to approximately 8%. In the second part of the second set, Del Potro won a game when Djokovic was serving (a so-called break) resulting in a large increase in Del Potro's WP. Although Del Potro won the second set, his WP did not increase by much, perhaps because the market had already anticipated that after a break, Del Potro would have managed to win the second set.

The outcome of the third and fourth sets was highly uncertain and was decided only after a tiebreak. Del Potro lost the third set, resulting in a large decrease in his WP from 33% (at the beginning of the tiebreak) to approximately 12%. In the middle of the fourth set, Djokovic won a game when Del Potro was serving (a break) resulting in a large decrease in Del Potro's WP, which reached a record low of 3%. Del Potro won the fourth set by a small margin, resulting in a spike in his winning probability from 4% to approximately 30%. Overall, this example makes it apparent that the betting prices readily react to new information and that the price adjustments are larger after tie-breaks because their outcome is more uncertain.

[22] In a live betting market, the price throughout the match of a betting asset, like "Del Potro to win the match", is represented by its in-play odds. By computing the inverse of these odds, we can derive the aggregate traders' belief about a given player's probability of winning the match at any moment of the match (Hasbrouck, 1991).

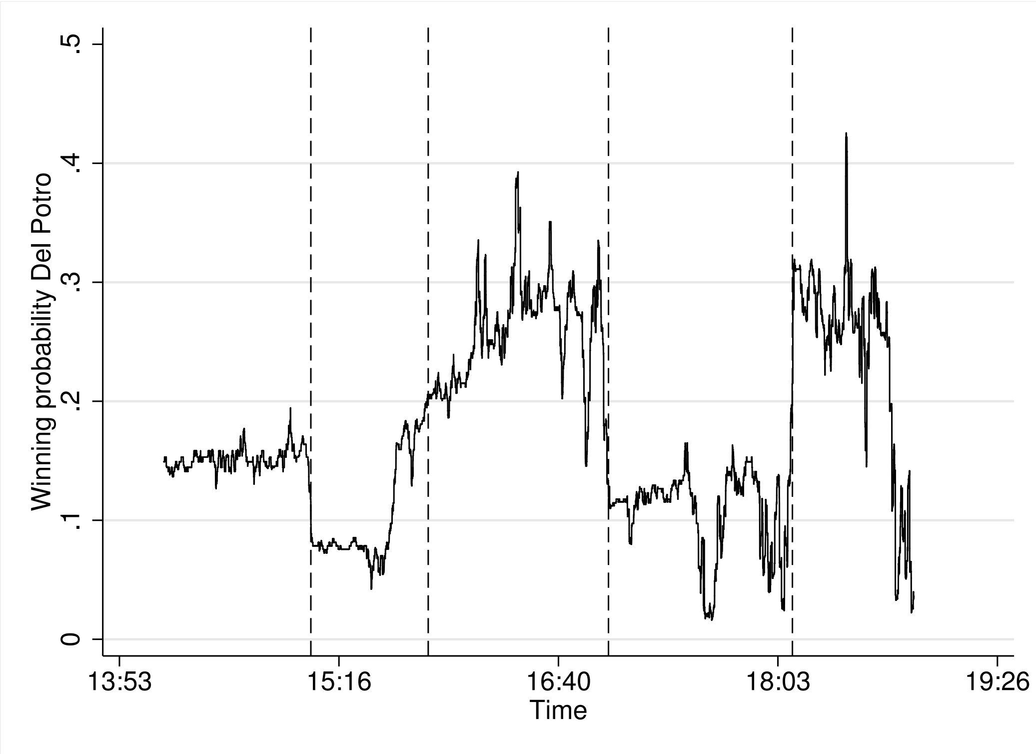


Figure 3
Displayed is Del Potro's odds-implied match winning probability for his 5th July, 2013 semifinal match against Djokovic at Wimbledon. Del Potro lost in five sets after more than four hours of play with the score: 5–7; 6–4; 7–6; 6–7; 3–6. A total of roughly \$74.8 million was bet on this match, 94% of which (\$70.5 million) was placed in-play.

5 Event study methods

The aim of our event study is to measure the betting odds reaction around the release of important information and to quantify the extent to which insider trading affects this reaction. Assuming rationality in the marketplace, security prices should immediately reflect all available information (Fama, 1970).

We follow the classical methodology presented by MacKinlay (1997). Fig. 4 summarizes the timeline of the baseline event study. We define the moment when a player wins a set or a tie-break as the event time (τ=0). We account for the possibility that some traders may anticipate the outcome of the point or that the umpire may update the score with a brief lag by letting our baseline event window begin two

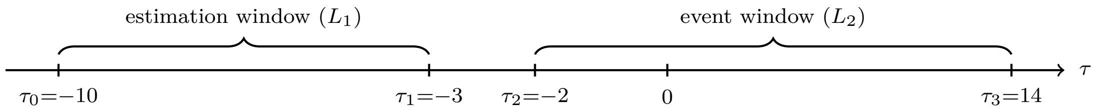


Figure 4
Displayed is the timeline of the baseline event study, where τ represents the time (in seconds) during a match. The news event takes place at τ=0. The estimation window length is L_1=8, and the event window length is L_2=17.

seconds before the event.[23] Thus, the event window spans from τ_1=−2 to τ_2=14 and lasts seventeen seconds (L_2=17).

The estimation window does not begin immediately after the end of the previous point (roughly at τ=−20) to avoid informational spillovers from the previous point. According to the ITF rules, the time between two consecutive points should be 20 seconds—in our sample, the median elapsed time is 22 seconds. Furthermore, as the event and estimation window should not overlap, the estimation window ends at τ_1=−3, one second before the start of the event window at τ_2=−2. Thus, the estimation window spans from τ_0=−10 to τ_1=−3 and lasts eight seconds (L_1=8).[24]

Because we want to analyze the reaction of the odds on the player who won the set, we consider only the odds of the set and tie-break winner. For each second τ of a match, we compute the average mid-odds from the best back ($back_\tau$) and sell odds (lay_τ) as follows:

$$mid_\tau = \frac{back_\tau + lay_\tau}{2} \quad . \tag{1}$$

The mid-price, which is expressed in decimal odds, can range from a minimum to 1.01 to a maximum of 1000. Then we derive the implied winning probability as follows by taking the inverse of the the mid-price ($WP_\tau = \frac{1}{mid_\tau}$), which approaches 99%

[23]The International Tennis Federation (ITF), the governing body of tennis, requires the umpires to “timely and accurately” enter the points in their computers throughout the match.

[24]In a robustness check, we also test an alternative estimation windows lasting 15 seconds, from τ_0=−17 to τ_1=−3 jointly with an alternative event windows lasting 19 seconds, from τ_2=−2 to τ_3=17. The results, which we present in Table A.3 in the Appendix, are similar.

when the mid-price approaches 1.01. Finally, we compute the actual returns from the odds-implied winning probability as follows:[25]

$$R_\tau = \frac{WP_\tau}{WP_{\tau-1}} - 1 \ . \tag{2}$$

When a player wins the set or tie-break, the odds should adjust downward, causing the odds-implied winning probability to increase and resulting in a positive return.

The drawback of any event study methodology is that the economic interpretation of the results clearly depends on the underlying assumptions used to estimate normal returns and for statistical testing. The null hypothesis that the average abnormal returns are zero implicitly includes a test of whether the model used for measuring normal returns is correct (MacKinlay, 1997). Since for tennis matches there is no asset valuation model to measure abnormal returns, the logical alternative is to use the mean return over the estimation window as an estimate for the normal return.[26]

We estimate the normal return by calculating the arithmetic mean of the returns over the estimation window for each event j in our sample as follows:

$$\bar{R}_j = \frac{1}{L_1} \sum_{\tau=\tau_0}^{\tau_1} R_{j\tau} \ , \tag{3}$$

where L_1 is the estimation window length. We then compute the abnormal returns (or prediction errors) for each event j by subtracting the normal returns from the actual returns over the event window: $AR_{j\tau} = R_{j\tau} - \bar{R}_{j\tau}$. Under the null hypothesis, the abnormal returns will be jointly normally distributed with a zero conditional mean and conditional variance $\sigma^2(AR_j)$. It can be assumed that $\sigma^2_{AR_j} = \sigma^2_{\epsilon_j}$, where $\sigma^2_{\epsilon_j}$

[25] Another option would be to compute the returns directly from the odds ($R_\tau = \frac{mid_\tau}{mid_{\tau-1}} - 1$). However, this approach results in negative returns when a player wins the set and thus offers a less intuitive interpretation.

[26] As a robustness check, we assume that the betting prices would not change and that the normal returns would be zero over the event window for all set events. The results are similar.

is estimated by computing the sample variance of the returns over the estimation window (MacKinlay, 1997) as follows:

$$\hat{\sigma}^2_{\epsilon_j} = \frac{1}{L_1 - 1} \sum_{\tau=\tau_0}^{\tau_1} (R_{j\tau} - \bar{R}_j)^2 \ . \tag{4}$$

Therefore, the null distribution of the abnormal returns is $AR_{j\tau} \sim N(0, \hat{\sigma}^2_{\epsilon_j})$.

Then, we aggregate the abnormal returns along two dimensions. First, we aggregate the returns along the event dimension to generate the average abnormal returns AAR_τ. To do so, for every second τ in the event window, we average the abnormal returns over the total number of events N: $AAR_\tau = \frac{1}{N}\sum_{j=1}^{N} AR_{j\tau}$. The variance of the AAR_τ is:

$$\hat{\sigma}^2_{AAR_\tau} = \frac{1}{N^2} \sum_{j=1}^{N} \hat{\sigma}^2_{\epsilon_j} \ . \tag{5}$$

Second, we aggregate along the time dimension the average abnormal returns to compute the cumulative average abnormal returns $CAAR_{(\tau_2,\tau_3)}$. We do so by summing up the average abnormal returns over the event window: $CAAR_{(\tau_2,\tau_3)} = \sum_{\tau=\tau_2}^{\tau_3} AAR_\tau$. The variance of the $CAAR_{(\tau_2,\tau_3)}$ is:

$$\hat{\sigma}^2_{CAAR_{(\tau_2,\tau_3)}} = \sum_{\tau=\tau_2}^{\tau_3} \hat{\sigma}^2(AAR_\tau) \ . \tag{6}$$

Under the assumption that returns are independent and identically normally distributed, the following test statistic for the average abnormal returns can be computed:

$$\theta_1 = \frac{AAR_\tau}{\hat{\sigma}_{AAR_\tau}} \overset{H_0}{\sim} N(0,1) \ , \tag{7}$$

whereas for the cumulative average abnormal returns, the test statistic is:

$$\theta_2 = \frac{CAAR_{(\tau_2,\tau_3)}}{\hat{\sigma}_{CAAR_{(\tau_2,\tau_3)}}} \overset{H_0}{\sim} N(0,1) \ . \tag{8}$$

6 Empirical results

6.1 Event study

Fig. 5 presents the cumulative average abnormal returns ($CAAR$) and the corresponding 95% confidence interval over the event window for the 365 set events.[27] We observe a significant positive reaction in the betting odds in the first seconds immediately following the set event. Most importantly, during the first five insider trading seconds the $CAAR$ is 3.06% and is statistically significant at the 1% level. Therefore, some sophisticated insider traders are able to circumvent the speed bump by timely canceling—without any delay—their limit orders on the wrong side of the market, as described in *Scenario 3*.

The pattern displayed in Fig. 5 shows that new information is fully incorporated into the betting odds within six to seven seconds, after which the $CAAR$ stabilizes around 5%. Thus, insider trading accounts for more than 60% of the full price reaction observed once the public receives the information about the set winner. This result shows that insider trading substantially contributes to "speeding up" the price discovery process, which contrasts with *Hypothesis 1.* This result is also in line with that of Meulbroek (1992), who finds, using illegal insider trading SEC data, that the price adjustment caused by insider traders accounts for between 40 and 50% of the total price adjustment.

Furthermore, Fig. 5 shows an average abnormal return (AAR) of 1.44%, the largest over the event window, at second six. Apparently, insider traders not only cancel the wrong outstanding limit orders but also immediately place new market orders (which undergo the five-second speed-bump) at the price that they deem correct to increase their profits. Finally, by slowing down insider traders, the speed bump somewhat lengthens the price discovery process by five seconds.

[27] In the Appendix, Panel A in Table A.2 reports the full results.

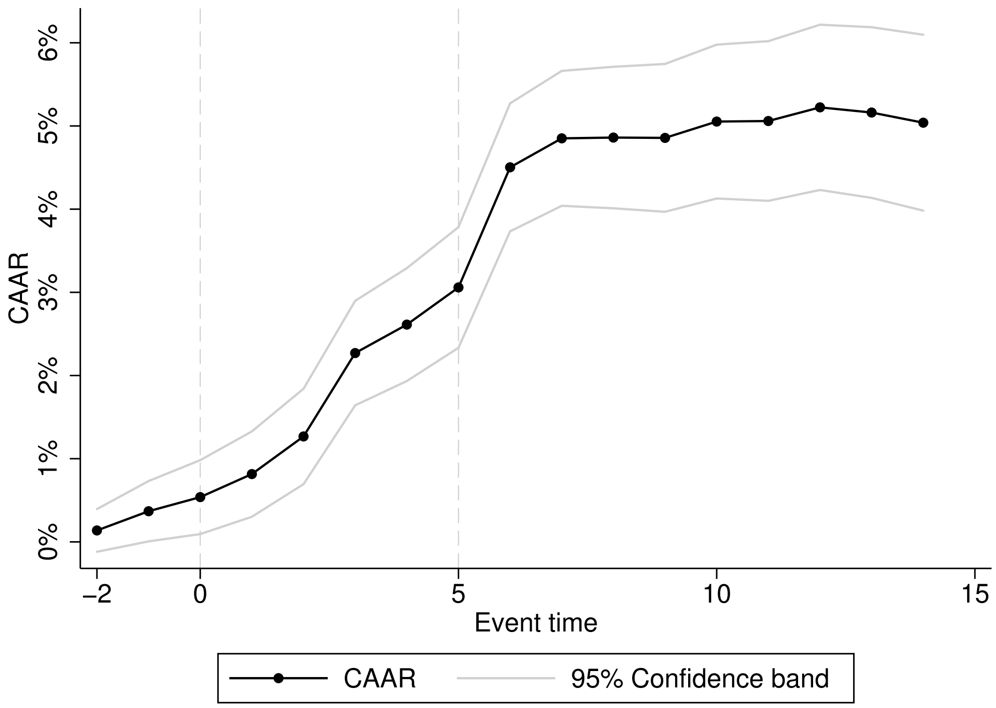


Figure 5
Displayed is the cumulative abnormal return averaged across the 365 event observations from second −2 to second 14. The event date (τ=0) is when a player wins the set. The realized returns are computed from the changes in the odds-implied winning probability of the player who won the set. The abnormal return is calculated using the constant mean return model.

Fig. 6 presents the cumulative average abnormal returns ($CAAR$) and the corresponding 95% confidence interval over the event window for the 79 tie-break events (i.e., sets won at the tie-break).[28] Overall, the $CAAR$ pattern is similar to that displayed in Fig. 5, with three main differences: first, the total adjustment of approximately 8.8% is significantly larger than for the set events (roughly 5%). This is due to the importance of winning a tie-break in a balanced match and the difficulty of predicting its outcome. Second, the largest average abnormal return (AAR) of 3.18% now takes place at τ=3, perhaps indicating that the insiders trade even faster or speculate to gain a edge over other traders during pivotal moments. Third, insider trading during the first five seconds causes around 80% of the full price reaction, which is larger than for the set events.

[28]In the Appendix, Panel B in Table A.2 reports the full results.

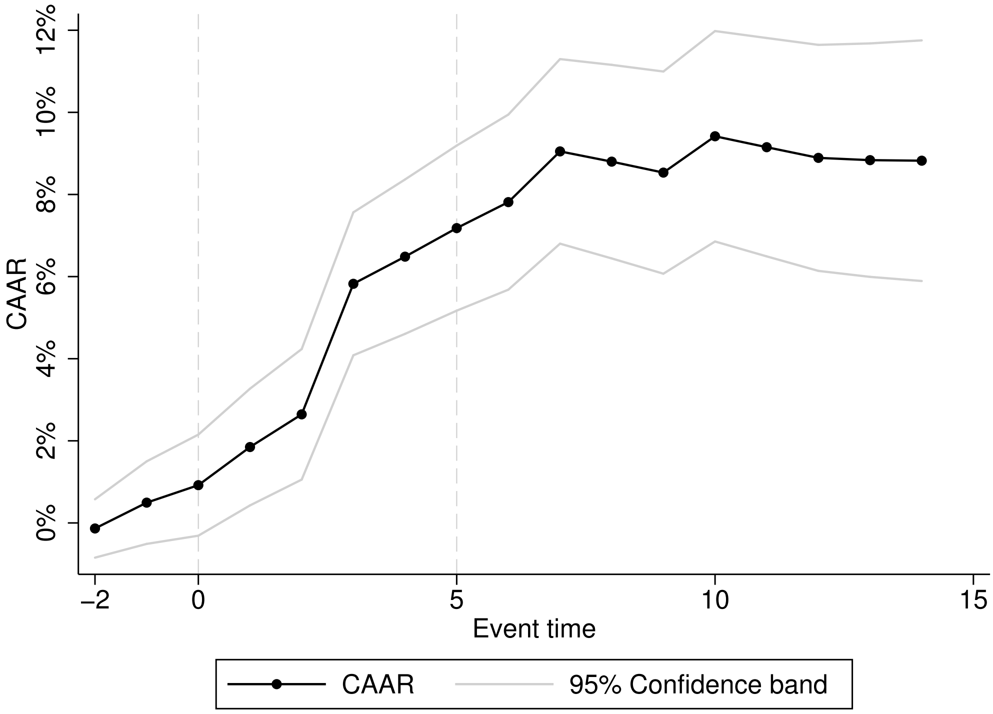


Figure 6
Displayed is the cumulative abnormal return averaged across the 79 event observations from second −2 to second 14. The event date (τ=0) is when a player wins the set in the tie-break. The realized returns are computed from the changes in the odds-implied winning probability of the player who won the tie-break. The abnormal return is calculated using the constant mean return model.

Taken together, both the results displayed in Fig. 5 and 6 show that the efficiency of price discovery is affected by insider trading, which contradicts *Hypothesis 1*. Furthermore, the tie-break results show that the impact of insider traders on price discovery is even larger during important moments, when inside information is more valuable.

6.2 Trading strategy

By knowing before others in which direction the betting odds will move, an insider may adopt a dynamic back-lay trading strategy to earn a profit by first backing the player who wins a set and later laying that same player. In order to earn a positive return, the trader must lay only when the lay odds have shortened and are lower than the previous back odds. By taking both sides, betting for and against the same player, the insider trader locks in a guaranteed profit regardless of the final outcome of the match.

More precisely, if Player 1 (P1) wins a set, the insider should back P1 at the higher (back) odds of 2.0 before the market moves. In the following seconds, the market reacts to the new information, and the odds for P1 decrease to reflect the increase in his probability of winning the match. If the lay odds are later lower than 2.0, say 1.5, the insider lays P1, thereby closing the trading strategy. By staking a certain amount on both sides of the market, the trader can lock in a risk-free profit, regardless of whether P1 wins or loses the match.[29]

As an illustration, consider the following strategy: backing $1 at odds of 2.0 just after a set event and laying $1.33 at odds of 1.50 some seconds later on P1. In a so-called balanced back-lay trading, the net return is 31% regardless of the outcome: if P1 wins the match, the back-bet payoff is $1 and the lay-bet payoff is $–0.67, yielding a gross profit of $0.33 (or a net profit of $0.31 after deducting the 5% Betfair commission); if P1 loses the match, the back-bet payoff is $–1 and the lay-bet payoff is $1.33, yielding exactly the same profit.[30]

We assess the total return that an insider trader could have earned by using the back–lay trading strategy when available for the set and tie-break events. Specifically, we assume that the insider trader backs the player who wins the set and lays the same player after exactly five seconds. For simplicity, we ignore of the speed bump and assume that the new market orders are immediately matched.

As we explained before, such a trading strategy yields a positive profit only if the back price is higher than the lay price and yields a larger profit when the back-lay difference is larger. This is more often the case in balanced sets (and matches) than in unbalanced ones. In an unbalanced set, when a player has a large advantage, the

[29] Hedging a trade by taking the opposite position on the same player is a strategy called "greening up". The name derives from the fact that by dynamically backing and laying the same player correctly, the trader makes a positive profit (indicated by a green number on the trading software) regardless of the outcome.

[30] The precise lay dollar amount yielding a balanced trade is computed by multiplying the back amount by the back odds and dividing by the lay odds.

odds usually move either by a small amount or not at all after the end of a set because the traders anticipate its outcome. To be the most profitable, this trading strategy is applied to balanced matches where the odds are around 1.5 and 2.0 so that the odds fluctuate more after an event.

We can identify 64 set (17.5% of all set events) and 31 tie-break events (39.2% of all tie-break events) that satisfy the condition in which the lay odds at second five are lower than the back odds at second zero. On average, we find than an insider using the dynamic back-lay trading strategy described above would earn a 5.32% return on set events and 5.97% on tie-break events (both returns are statistically significant at the 1% level). By putting \$100 at stake, the insider could earn between \$5 and \$6 without taking any risk.

In a further analysis, we consider only the trading opportunities during the semi-finals and the finals. We identify 22 trading opportunities for the set events, yielding 6.78%, and eight trading opportunities for the tie-break events, yielding 6.41%. As expected, the trading profits are larger because these matches are usually more balanced and the liquidity is higher.

Overall, these results are important as they show that in such situations insiders can make significant profits without taking any risk. Moreover, insiders can make even larger profits by, for example, waiting more than five seconds to close the trade and lay at even lower odds, or by taking some risks by stacking a larger amount on the back-bet than on the lay-bet on the winning player (an unbalanced trade). However, one should not forget that insiders not only have to cover their travel, stay, and entrance costs but also have to deal with the stadium security agents.

7 Concluding remarks

In most financial markets, trading on material information by corporate insiders is against the law. Thus, the evidence on the extent to which this important form of trading aids the price discovery is, at best, scarce. We attempt to fill this gap in the literature by using a previously unexplored setting: in-play tennis betting. Due to inevitable communication delays, betting traders in the stadium have an informational advantage because they can observe material information about the match before other traders, such as TV traders. Thus, this setting offers an unique opportunity to distinguish the in-play trading activity of insiders from that of outsiders.

Our event study analyses of 365 set events from 141 men's singles matches at two major professional tennis tournaments show that the betting prices rapidly adjust in the first five insider trading seconds following important moments—the end of sets and tie-breaks—when outsiders have not yet received the information. Most importantly, we show that the cumulative abnormal return during the insider trading seconds following an important event is between 60% and 80% of the full price reaction observed once the public receives the new information. Thus, the price discovery is largely affected by insider trading. We also show that the the impact of insider trading is even larger for unanticipated news events, like tie-break set events, when inside information is more valuable. Finally, we estimate that a simple dynamic back-lay strategy implemented in the first seconds after the event yields large risk-free trading profits to insiders ranging between 5% and 7%.

Our results are important, as they provide empirical evidence on the advantages of insider trading in terms of its contribution to price efficiency. Insider traders have large incentives to rapidly integrate the new information into the market. At the same time, as betting on a betting exchange is a zero-sum game, one could also discuss how the insider activity disadvantages the slower traders. Although Betfair has put a speed bump in place in an effort to protect these slower traders, our data suggest that sophisticated insiders have developed strategies to circumvent it. Further research should determine if slower traders adopt more passive trading strategies or even stop trading when they anticipate that they cannot compete with insiders on speed. Overall, when debating the effects of insider trading on a betting exchange as well as on any financial market, these adverse selection costs to slower traders should be weighed against the positive externalities of greater price efficiency.

A Appendix

A.1 Tennis

This appendix introduces the basic rules of tennis and its jargon. These rules are specific to Grand Slam tournaments and can be found on the website of the International Tennis Federation. At the beginning of the match, a coin toss decides which player starts serving in the first game. Player 1 begins the match by serving in the first game of the first set—player 2 is the receiver. A player wins a point, sometimes referred to as point game, if the opponent cannot return the ball. A game is won when one player wins four points with a two-point difference, or when there is a two-point difference after a deuce, i.e., a score of 40–40 (3 points to 3 points in a game). A player has a gamepoint if he needs one more point to win the game: if this player is the receiver, the situation is called a breakpoint. A break (of service) happens when the receiver wins the game.

The players alternate serving every game, and they change ends after every odd-numbered game. A *set* is won when a player either wins six games with a two-game difference, or, in the case of a tie-break, when the score for one player is 7:6. A player has a setpoint if he needs one more point to win a set: if this player is the receiver, he has a breakpoint to win the set.

The *tie-break* begins when the game score is tied at 6:6 and is played until one player wins seven points with a two-point difference, or until there is a two-point difference when the point score is 6:6. At Grand Slam tournaments, a tennis match is played as the best of five sets, meaning that the first player winning three sets wins the match. The fifth set does not have a tie-break; the set (and match) is won when one player wins two more games than the other player.

Table A.1
List of the matches included in our sample.

Player 1	Player 2	Date	Stage	Player 1	Player 2	Date	Stage
French Open:				**Wimbledon (cont.):**			
R.Gasquet	R.Stepanek	23-May-2011	1	A.Roddick	A.Murray	3-Jul-2009	6
A.Clement	M.Berrer	26-May-2011	2	A.Roddick	R.Federer	5-Jul-2009	7
J.Tipsarevic	R.Federer	27-May-2011	3	R.Federer	A.Falla	21-Jun-2010	1
S.Darcis	G.Monfils	27-May-2011	3	R.Federer	I.Bozoljac	23-Jun-2010	2
R.Gasquet	N.Djokovic	29-May-2011	4	R.Federer	A.Clement	25-Jun-2010	3
G.Simon	R.Soderling	30-May-2011	4	P.Petzschner	R.Nadal	26-Jun-2010	3
G.Monfils	R.Federer	31-May-2011	5	N.Djokovic	L.Hewitt	28-Jun-2010	4
A.Murray	J.Chela	1-Jun-2011	5	R.Federer	J.Melzer	28-Jun-2010	4
R.Nadal	R.Soderling	1-Jun-2011	5	R.Federer	T.Berdych	30-Jun-2010	5
R.Federer	N.Djokovic	3-Jun-2011	6	R.Soderling	R.Nadal	30-Jun-2010	5
R.Nadal	A.Murray	3-Jun-2011	6	J.Tsonga	A.Murray	30-Jun-2010	5
R.Nadal	R.Federer	5-Jun-2011	7	A.Murray	R.Nadal	2-Jul-2010	6
M.Llodra	G.Garcia-Lopez	28-May-2012	1	T.Berdych	N.Djokovic	2-Jul-2010	6
I.Sijsling	G.Muller	27-May-2012	1	T.Berdych	R.Nadal	4-Jul-2010	7
M.Berrer	J.Melzer	27-May-2012	1	M.Kukushkin	R.Federer	21-Jun-2011	1
R.Federer	A.Ungur	30-May-2012	2	A.Mannarino	R.Federer	23-Jun-2011	2
J.Del Potro	M.Cilic	1-Jun-2012	3	M.Baghdatis	N.Djokovic	25-Jun-2011	3
F.Fognini	Jw.Tsonga	1-Jun-2012	3	D.Nalbandian	R.Federer	25-Jun-2011	3
G.Simon	S.Wawrinka	1-Jun-2012	3	A.Murray	R.Gasquet	27-Jun-2011	4
J.Monaco	R.Nadal	4-Jun-2012	4	M.Llodra	N.Djokovic	27-Jun-2011	4
J.Tipsarevic	N.Almagro	4-Jun-2012	4	M.Youzhny	R.Federer	27-Jun-2011	4
R.Federer	D.Goffin	3-Jun-2012	4	R.Nadal	M.Fish	29-Jun-2011	5
N.Djokovic	A.Seppi	3-Jun-2012	4	J.Tsonga	R.Federer	29-Jun-2011	5
N.Almagro	R.Nadal	6-Jun-2012	5	R.Nadal	A.Murray	1-Jul-2011	6
D.Ferrer	A.Murray	6-Jun-2012	5	J.Tsonga	N.Djokovic	1-Jul-2011	6
R.Federer	J.Del Potro	5-Jun-2012	5	R.Nadal	N.Djokovic	3-Jul-2011	7
N.Djokovic	J.Tsonga	5-Jun-2012	5	R.Federer	A.Ramos	25-Jun-2012	1
D.Ferrer	R.Nadal	8-Jun-2012	6	R.Federer	F.Fognini	27-Jun-2012	2
N.Djokovic	R.Federer	8-Jun-2012	6	R.Federer	J.Benneteau	29-Jun-2012	3
M.Raonic	M.Llodra	29-May-2013	2	R.Federer	X.Malisse	2-Jul-2012	4
G.Monfils	E.Gulbis	29-May-2013	2	R.Federer	M.Youzhny	4-Jul-2012	5
M.Przysiezny	R.Gasquet	31-May-2013	2	D.Ferrer	A.Murray	4-Jul-2012	5
N.Djokovic	G.Pella	30-May-2013	2	A.Murray	J.Tsonga	6-Jul-2012	6
V.Troicki	M.Cilic	31-May-2013	3	N.Djokovic	R.Federer	6-Jul-2012	6
N.Davydenko	R.Gasquet	1-Jun-2013	3	R.Federer	A.Murray	8-Jul-2012	7
T.Haas	J.Isner	1-Jun-2013	3	V.Hanescu	R.Federer	24-Jun-2013	1
K.Anderson	D.Ferrer	2-Jun-2013	4	R.Nadal	S.Darcis	24-Jun-2013	1
R.Nadal	K.Nishikori	3-Jun-2013	4	N.Djokovic	F.Mayer	25-Jun-2013	1
J.Tsonga	R.Federer	4-Jun-2013	5	D.Ferrer	M.Alund	25-Jun-2013	1
T.Robredo	D.Ferrer	4-Jun-2013	5	A.Ramos	J.Del Potro	25-Jun-2013	1
R.Nadal	S.Wawrinka	5-Jun-2013	5	J.Tsonga	E.Gulbis	26-Jun-2013	2
N.Djokovic	T.Haas	5-Jun-2013	5	F.Verdasco	J.Benneteau	26-Jun-2013	2
D.Ferrer	Jw.Tsonga	7-Jun-2013	6	N.Djokovic	B.Reynolds	27-Jun-2013	2
N.Djokovic	R.Nadal	7-Jun-2013	6	R.Gasquet	G.Soeda	27-Jun-2013	2
R.Nadal	D.Ferrer	9-Jun-2013	7	J.Levine	J.Del Potro	27-Jun-2013	2
V.Estrella Burgos	J.Janowicz	25-May-2014	1	J.Melzer	S.Stakhovsky	28-Jun-2013	3
A.Dolgopolov	A.Ramos	25-May-2014	1	J.Janowicz	J.Melzer	1-Jul-2013	4
G.Elias	D.Schwartzman	25-May-2014	1	M.Youzhny	A.Murray	1-Jul-2013	4
R.Gasquet	B.Tomic	27-May-2014	1	N.Djokovic	T.Berdych	3-Jul-2013	5
G.Monfils	V.Hanescu	27-May-2014	1	F.Verdasco	A.Murray	3-Jul-2013	5
I.Karlovic	K.Anderson	31-May-2014	3	N.Djokovic	J.Del Potro	5-Jul-2013	6
E.Gulbis	R.Federer	1-Jun-2014	4	J.Janowicz	A.Murray	5-Jul-2013	6
T.Berdych	J.Isner	1-Jun-2014	4	N.Djokovic	A.Murray	7-Jul-2013	7
G.Garcia-Lopez	G.Monfils	2-Jun-2014	4	G.Dimitrov	R.Harrison	23-Jun-2014	1
M.Raonic	N.Djokovic	3-Jun-2014	5	A.Murray	D.Goffin	23-Jun-2014	1
T.Berdych	E.Gulbis	3-Jun-2014	5	P.Lorenzi	R.Federer	24-Jun-2014	1
G.Monfils	A.Murray	4-Jun-2014	5	S.Wawrinka	J.Sousa	24-Jun-2014	1
R.Nadal	D.Ferrer	4-Jun-2014	5	G.Dimitrov	L.Saville	25-Jun-2014	2
E.Gulbis	N.Djokovic	6-Jun-2014	6	A.Murray	B.Rola	25-Jun-2014	2
R.Nadal	A.Murray	6-Jun-2014	6	G.Muller	R.Federer	26-Jun-2014	2
R.Nadal	N.Djokovic	8-Jun-2014	7	L.Rosol	R.Nadal	26-Jun-2014	2
Wimbledon:				S.Wawrinka	Y-H.Lu	26-Jun-2014	2
G.Garcia-Lopez	R.Federer	24-Jun-2009	2	N.Djokovic	G.Simon	27-Jun-2014	3
P.Kohlschreiber	R.Federer	26-Jun-2009	3	S.Giraldo	R.Federer	28-Jun-2014	3
S.Wawrinka	J.Levine	27-Jun-2009	3	M.Kukushkin	R.Nadal	28-Jun-2014	3
A.Murray	S.Wawrinka	29-Jun-2009	4	S.Wawrinka	D.Istomin	30-Jun-2014	3
R.Soderling	R.Federer	29-Jun-2009	4	A.Murray	G.Dimitrov	2-Jul-2014	5
T.Haas	N.Djokovic	1-Jul-2009	5	S.Wawrinka	R.Federer	2-Jul-2014	5
L.Hewitt	A.Roddick	1-Jul-2009	5	N.Djokovic	G.Dimitrov	4-Jul-2014	6
I.Karlovic	R.Federer	1-Jul-2009	5	R.Federer	M.Raonic	4-Jul-2014	6
T.Haas	R.Federer	3-Jul-2009	6	N.Djokovic	R.Federer	6-Jul-2014	7

Notes: The table lists all 141 matches played at the French Open (Roland Garros) and at the Wimbledon Championships included in our sample. Stage indicates the tournament stage, from a first-round match (Stage=1) up to the final match (Stage=7).

Table A.2
Event study results (baseline).

Panel A: set events (N=365)

Time (τ)	AAR	t-statistic (θ_1)	p-value	$CAAR$	t-statistic (θ_2)	p-value
–2	0.14%	1.05	0.293	0.14%	1.05	0.293
–1	0.23%	1.76	0.077	0.37%	1.99	0.046
0	0.17%	1.29	0.194	0.54%	2.37	0.017
1	0.28%	2.11	0.034	0.82%	3.11	0.001
2	0.45%	3.45	0.000	1.27%	4.33	0.000
3	1.00%	7.66	0.000	2.27%	7.08	0.000
4	0.34%	2.61	0.008	2.61%	7.54	0.000
5	0.45%	3.41	0.000	3.06%	8.26	0.000
6	1.44%	11.03	0.000	4.50%	11.46	0.000
7	0.35%	2.66	0.007	4.85%	11.72	0.000
8	0.01%	0.07	0.941	4.86%	11.19	0.000
9	0.00%	-0.02	1.023	4.86%	10.71	0.000
10	0.20%	1.49	0.134	5.05%	10.7	0.000
11	0.01%	0.05	0.958	5.06%	10.33	0.000
12	0.16%	1.25	0.209	5.22%	10.3	0.000
13	-0.06%	-0.47	1.365	5.16%	9.86	0.000
14	-0.12%	-0.93	1.649	5.04%	9.33	0.000

Panel B: tie-break events (N=79)

Time (τ)	AAR	t-statistic (θ_1)	p-value	$CAAR$	t-statistic (θ_2)	p-value
–2	-0.13%	-0.37	1.289	-0.13%	-0.37	1.289
–1	0.63%	0.97	0.082	0.49%	0.97	0.334
0	0.43%	1.47	0.24	0.92%	1.47	0.143
1	0.93%	2.55	0.01	1.85%	2.55	0.011
2	0.80%	3.26	0.028	2.65%	3.26	0.001
3	3.18%	6.56	0.000	5.82%	6.56	0.000
4	0.66%	6.76	0.068	6.48%	6.76	0.000
5	0.70%	7.00	0.054	7.18%	7.00	0.000
6	0.63%	7.18	0.081	7.81%	7.18	0.000
7	1.24%	7.89	0.000	9.05%	7.89	0.000
8	-0.25%	7.32	1.509	8.80%	7.32	0.000
9	-0.27%	6.79	1.537	8.53%	6.79	0.000
10	0.89%	7.20	0.014	9.42%	7.20	0.000
11	-0.27%	6.74	1.538	9.15%	6.74	0.000
12	-0.26%	6.33	1.526	8.89%	6.33	0.000
13	-0.06%	6.09	1.121	8.84%	6.09	0.000
14	-0.01%	5.90	1.028	8.82%	5.90	0.000

Notes: The table reports the average abnormal returns (AAR) and the cumulative average abnormal returns ($CAAR$) over the event window. In the baseline analysis, the event window spans from τ_1=−2 to τ_2=14 and lasts 17 seconds (L_2=17) whereas the estimation window spans from τ_0=−10 to τ_1=−3 and lasts eight seconds (L_1=8).

Table A.3
Event study results (robustness check).

Panel A: set events (N=365)						
Time (τ)	AAR	t-statistic (θ_1)	p-value	$CAAR$	t-statistic (θ_2)	p-value
–2	0.09%	0.29	0.775	0.09%	0.29	0.775
–1	0.19%	0.57	0.567	0.28%	0.61	0.544
0	0.12%	0.38	0.700	0.40%	0.72	0.473
1	0.23%	0.72	0.474	0.63%	0.98	0.327
2	0.41%	1.26	0.209	1.04%	1.44	0.150
3	0.96%	2.96	0.003	2.00%	2.52	0.012
4	0.30%	0.92	0.359	2.30%	2.68	0.007
5	0.40%	1.24	0.215	2.70%	2.94	0.003
6	1.40%	4.32	0.000	4.10%	4.22	0.000
7	0.30%	0.94	0.349	4.40%	4.30	0.000
8	-0.04%	-0.11	1.088	4.36%	4.06	0.000
9	-0.05%	-0.15	1.120	4.32%	3.85	0.000
10	0.15%	0.47	0.641	4.47%	3.82	0.000
11	-0.04%	-0.12	1.094	4.43%	3.65	0.000
12	0.12%	0.37	0.713	4.55%	3.62	0.000
13	-0.11%	-0.33	1.260	4.44%	3.43	0.001
14	-0.17%	-0.52	1.395	4.27%	3.20	0.001
15	1.06%	3.26	0.001	5.33%	3.88	0.000
16	-0.09%	-0.29	1.225	5.24%	3.71	0.000
17	-0.11%	-0.33	1.255	5.13%	3.54	0.000
Panel B: tie-break events (N=79)						
Time (τ)	AAR	t-statistic (θ_1)	p-value	$CAAR$	t-statistic (θ_2)	p-value
–2	0.06%	0.16	0.876	0.06%	0.16	0.876
–1	0.83%	2.04	0.042	0.89%	1.55	0.121
0	0.62%	1.54	0.124	1.51%	2.15	0.031
1	1.13%	2.78	0.006	2.64%	3.25	0.001
2	0.99%	2.45	0.014	3.63%	4.01	0.000
3	3.38%	8.32	0.000	7.01%	7.05	0.000
4	0.86%	2.12	0.034	7.87%	7.33	0.000
5	0.89%	2.20	0.028	8.76%	7.64	0.000
6	0.83%	2.05	0.041	9.59%	7.88	0.000
7	1.43%	3.53	0.000	11.03%	8.59	0.000
8	-0.05%	-0.13	1.103	10.97%	8.16	0.000
9	-0.07%	-0.17	1.135	10.91%	7.76	0.000
10	1.08%	2.67	0.008	11.99%	8.19	0.000
11	-0.07%	-0.17	1.135	11.92%	7.85	0.000
12	-0.06%	-0.15	1.122	11.86%	7.55	0.000
13	0.14%	0.35	0.726	12.00%	7.39	0.000
14	0.18%	0.45	0.649	12.18%	7.28	0.000
15	0.15%	12.19	0.000	12.33%	7.95	0.000
16	-0.14%	-0.34	1.267	12.19%	7.61	0.000
17	-0.10%	-0.24	1.188	12.09%	7.31	0.000

Notes: The table reports the average abnormal returns (AAR) and the cumulative average abnormal returns ($CAAR$) over the event window. In the robustness analysis, the event window spans from τ_1=−2 to τ_2=17 and lasts 19 seconds (L_2 = 19) whereas the estimation window spans from τ_0=−17 to τ_1=−3 and lasts 15 seconds (L_1=15).

References

Betfair. (2015). *Annual report.* Retrieved 2017-06-10, from www.paddypowerbetfair.com/~/media/Files/P/Paddy-Power-Betfair/archive/annual-report-general-meeting/betfair/annual-report-2015.pdf

Brown, A. (2012). Evidence of in-play insider trading on a UK betting exchange. *Applied Economics*, *44*(9), 1169–1175.

Brown, A., & Yang, F. (2017). Slowing down fast traders: Evidence from the Betfair speed bump (forthcoming). *Journal of Financial Markets.*

Carlton, D. W., & Fischel, D. R. (1983). The regulation of insider trading. *Stanford Law Review*, 857–895.

Chakravarty, S., & McConnell, J. J. (1997). An analysis of prices, bid/ask spreads, and bid and ask depths surrounding Ivan Boesky's illegal trading in Carnation's stock. *Financial Management*, 18–34.

Chakravarty, S., & McConnell, J. J. (1999). Does insider trading really move stock prices? *Journal of Financial and Quantitative Analysis*, *34*(2), 191–209.

Choi, D., & Hui, S. K. (2014). The role of surprise: Understanding overreaction and underreaction to unanticipated events using in-play soccer betting market. *Journal of Economic Behavior & Organization*, *107*(Part B), 614–629.

Cornell, B., & Sirri, E. R. (1992). The reaction of investors and stock prices to insider trading. *The Journal of Finance*, *47*(3), 1031–1059.

Cox, S. (2015). *Why tennis courtsiding was my dream job.* Retrieved 2017-06-16, from www.bbc.com/news/magazine-32402945

Croxson, K., & Reade, J. (2014). Information and efficiency: Goal arrival in soccer betting. *The Economic Journal*, *124*(575), 62–91.

Dickson, C. (2015). Courtsiding in sport: Cheating, sharp practice or merely irritating? *Sports Law eJournal.*

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383–417.

Fama, E. F. (1991). Efficient capital markets 2. *The Journal of Finance*, *46*(5), 1575–1617.

Fernandes, N., & Ferreira, M. A. (2009). Insider trading laws and stock price informativeness. *Review of Financial Studies*, *22*(5), 1845–1887.

Fishman, M. J., & Hagerty, K. M. (1992). Insider trading and the efficiency of stock prices. *The RAND Journal of Economics*, *23*(1), 106–122.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *The Journal of Finance*, *46*(1), 179–207.

Hutchins, B. (2014). *Game, Set, Cash! Inside the secret world of international tennis trading.* Nero.

Ikenberry, D., Lakonishok, J., & Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics*, *39*(2), 181–208.

Jaffe, J. F. (1974). Special information and insider trading. *The Journal of Business*, *47*(3), 410–428.

Ke, B., Huddart, S., & Petroni, K. (2003). What insiders know about future earnings and how they use it: Evidence from insider trades. *Journal of Accounting and Economics*, *35*(3), 315–346.

Keown, A. J., & Pinkerton, J. M. (1981). Merger announcements and insider trading activity: An empirical investigation. *The Journal of Finance*, *36*(4), 855–869.

Kooij, W. J., Stokking, H. M., van Brandenburg, R., & de Boer, P.-T. (2014). Playout delay of TV signals: Measurement system design, validation and results. In *Proceedings of the 2014 ACM international conference on interactive experiences for TV and online video* (pp. 23–30).

Lakonishok, J., & Lee, I. (2001). Are insider trades informative? *Review of Financial Studies*, *14*(1), 79–111.

Leland, H. E. (1992). Insider trading: Should it be prohibited? *Journal of Political Economy*, *100*(4), 859–887.

MacKinlay, A. C. (1997). Event studies in economics and finance. *Journal of Economic Literature*, *35*(1), 13–39.

Manne, H. G. (1966). Defense of insider trading. *Harvard Business Review*, *44*(6), 113–122.

Meulbroek, L. K. (1992). An empirical analysis of illegal insider trading. *The Journal of Finance*, *47*(5), 1661–1699.

Mitchell, M., Pulvino, T., & Stafford, E. (2004). Price pressure around mergers. *The Journal of Finance*, *59*(1), 31–63.

Rozeff, M. S., & Zaman, M. A. (1988). Market efficiency and insider trading: New evidence. *Journal of Business*, *61*(1), 25–44.

Schnytzer, A., & Shilony, Y. (1995). Inside information in a betting market. *The Economic Journal*, *105*(431), 963–971.

Townend, M. (2016). Tennis gambling market second only to football for bookmakers after boom in online and in-play betting. *Daily Mail Online*.

Williams, L. V. (2005). *Information efficiency in financial and betting markets.* Cambridge University Press.

Curriculum vitae

Personal information

Paolo Bizzozero	born on 18.09.1987

Education

Ph.D. studies in Management and Economics University of Zurich	11/2013 – 10/2017
M.Sc. in Economics and Management Sciences Humboldt University of Berlin	10/2010 – 9/2013
B.Sc. in Economics HEC University of Lausanne	10/2006 – 6/2009

Professional experience

Teaching and research associate University of Zurich	10/2013 – present
Intern, Investment Product and Services UBS AG	9/2012 – 2/2013

Other

Teaching Skills certificate University of Zurich	1/2014 – 2/2017